

## JOURNAL

Submitted (03-DEC-2001) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA

## REMARK

NIH-MGC Project URL: <http://mgc.nci.nih.gov>  
Contact: MGC help desk  
Email: [cgaps-remail.nih.gov](mailto:cgaps-remail.nih.gov)  
Tissue Procurement: ATCC  
cDNA Library Preparation: Life Technologies, Inc.  
DNA Sequencing by: Sequencing Group at the Stanford Human Genome Center, Stanford University School of Medicine, Stanford, CA 94305  
Web site: <http://www-shgc.stanford.edu>  
Contact: (Dickson, Mark) [mdcpaxil.stanford.edu](mailto:mdcpaxil.stanford.edu)  
Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers, R. M.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LINL at: <http://image.llnl.gov>  
Series: IRAX Plate: 20 Row: 1 Column: 3  
This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 16507953.  
Location/Qualifiers  
1..2621

## FEATURES

## source

## CDS

/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="MGC:9539 IMAGE:3847282"  
/tissue\_type="Colon, adenocarcinoma"  
/clone\_id="NIH MGC\_65"  
/lab\_host="DH10B"  
/note="vector: pCMV-SPORT6"  
178..2400  
/codon\_start=1  
/product="similar to caspase recruitment domain protein 14"  
/protein\_id="AAH18142.1"  
/db\_xref="gi:17390315"  
/translation="MGEICRRDSALTALDEETLWEMESHRRHRIYRCICPSRLTPYLQAKVLCQIDEEVYLSPLTNSARAGHLIDLKTRGNKAIFALSLKFNHNDVTLVTGLQPDVDFNSFGLMETSKLTETCLAGISLOEINOEKQKQVILRRCCQLOEHLGLAETBAGLHQLADHSRMKREVSAAHFHEVILRLKDEMLSLSHYSNALOEKLAASCRSLQEBYLILKQELQKQANVSSCELELQEOSLRTASDQSGEBELNRLKENEXKASLTFSLAEKDLIEQSLDEARSGROELVERIHSLEKRAVAARQREQWEEKQVLLPQKSNACQLYREKVAQAQVCELOKEDQAVSARQREISQSLVEKQSLRQVLELTQVCELTQRLQALPQVILKQEARTRPCREKQRLRMAHLCPRDSCEIIVSTEQSLDLSATSRRELVDSPRSSPAPQOQLYKRAVDPEBEPSPSCIEIBEGDQGLPQAKGDPHLIDYEILDTDLPOLSESLQPSRGLDVSEGVIMRRPARLISQVTMLAFQGDALILEQISVIGANTLGIPIHRTVTPGSAADOMALRPGQIVMDYEASEPLPRAVLIEDTLLEAVGLILRRVDGFCLSVKNTDGYKRLLOLEAKVAISGSPFYIRVNLMEGRAKGELQVHCNEVILHTVDTMFGQGCWHARVNSYTMKDTAAHGTTIPNYSR"

## BASE COUNT

563 a 782 c 857 g 419 t

Query Match 84.2%; Score 351; DB 9; Length 2621;  
Best Local Similarity 100.0%; Pred. No. 9.1e-58;  
Matches 351; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 ATGGGGGAAGTGTGTCGGGAGGAGTCCGACATCACGGGACATGACGAGAGACACTGTG 60  
178 ATGGGGGAAGTGTGTCGGGAGGAGTCCGACATCACGGGACATGACGAGAGACACTGTG 237  
61 GAGATGATGAGAGACCGGACAGATGATGATGATGATGATGATGATGATGATGATGATG 120  
238 GAGATGATGAGAGACCGGACAGATGATGATGATGATGATGATGATGATGATGATGATG 297  
121 CCTTACCTGGCCGAGGACCAAGTGTGTCGACGATGACGAGAGAGAGAGTGTCTCAACG 180  
298 CCTTACCTGGCCGAGGACCAAGTGTGTCGACGATGACGAGAGAGAGAGTGTCTCAACG 357  
181 CCCCAGCTACCAACAGCGCATGCGGCGCGGCGCATCTTGCTGATTGCTGAAGACTCGA 240

## Db

358 CCCCAGCTACCAACAGCGCATGCGGCGCGGCGCATCTTGCTGATTGCTGAAGACTCGA 417

## Qy

241 GGGAGAAAGCGGGGCGCATGCTTCTTGAGAGACCTGAAGTTCACACACCTGACCTTAC 300

## Db

418 GGGAGAAAGCGGGGCGCATGCTTCTTGAGAGACCTGAAGTTCACACACCTGACCTTAC 477

## Qy

301 ACCCTGGTACCGGGCTGACGCTGATGATGATGATGATGATGATGATGATGATGATGAT 351

## Db

478 ACCCTGGTACCGGGCTGACGCTGATGATGATGATGATGATGATGATGATGATGATGAT 528

## RESULT 2

AY032927 3682 bp mRNA linear PRI 07-DEC-2001  
AY032927  
LOCUS  
DEFINITION  
complete cds.

## ACCESSION

AY032927.1 GI:14192722

## VERSION

AY032927

## KEYWORDS

complete cds.

## SOURCE

Homo sapiens (human)

## ORGANISM

Homo sapiens (human)

## REFERENCE

Gaido, O., Martinon, F., Micheau, O., Bonnet, D., Thome, M., and Tschopp, J.

## AUTHORS

Carnal, a CARD-containing binding partner of Bcl10, induces Bcl10 phosphorylation and NF-kappaB activation

## TITLE

FEBS Lett. 496 (2-3), 121-127 (2001)

## JOURNAL

21255663

## PUBMED

11356195

## REFERENCE

2 (bases 1 to 3682)

## AUTHORS

Martinon, F., Thome, M., and Tschopp, J.

## TITLE

Direct Submision

## JOURNAL

Submitted (23-Apr-2001) Institute of Biochemistry, University of Lausanne, Ch. des Boveresses 155, Epalinges, VD 1066, Switzerland

## FEATURES

Location/Qualifiers  
1..3682  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
1..3682  
/gene="CARMA2"  
4..3018  
/note="contains caspase recruitment domain CARD, coiled coil, PDZ, SH3, guanylate kinase domain GUK, interacts with the CARD-containing molecule Bcl10 (CARDEN)"  
/codon\_start=1  
/product="CARD-containing MAGUK 2 protein"  
/protein\_id="AAK54453.1"  
/db\_xref="gi:14192723"

## translation

"MGEICRRDSALTALDEETLWEMESHRRHRIYRCICPSRLTPYLQAKVLCQIDEEVYLSPLTNSARAGHLIDLKTRGNKAIFALSLKFNHNDVTLVTGLQPDVDFNSFGLMETSKLTETCLAGISLOEINOEKQKQVILRRCCQLOEHLGLAETBAGLHQLADHSRMKREVSAAHFHEVILRLKDEMLSLSHYSNALOEKLAASCRSLQEBYLILKQELQKQANVSSCELELQEOSLRTASDQSGEBELNRLKENEXKASLTFSLAEKDLIEQSLDEARSGROELVERIHSLEKRAVAARQREQWEEKQVLLPQKSNACQLYREKVAQAQVCELOKEDQAVSARQREISQSLVEKQSLRQVLELTQVCELTQRLQALPQVILKQEARTRPCREKQRLRMAHLCPRDSCEIIVSTEQSLDLSATSRRELVDSPRSSPAPQOQLYKRAVDPEBEPSPSCIEIBEGDQGLPQAKGDPHLIDYEILDTDLPOLSESLQPSRGLDVSEGVIMRRPARLISQVTMLAFQGDALILEQISVIGANTLGIPIHRTVTPGSAADOMALRPGQIVMDYEASEPLPRAVLIEDTLLEAVGLILRRVDGFCLSVKNTDGYKRLLOLEAKVAISGSPFYIRVNLMEGRAKGELQVHCNEVILHTVDTMFGQGCWHARVNSYTMKDTAAHGTTIPNYSR"

## BASE COUNT

751 a 1100 c 1194 g 637 t

## ORIGIN

1100 c 1194 g 637 t



```

/product="caspase recruitment domain protein 14"
/db_xref="GI:12382775"
/translation="MGELCRDPSALMTALDEETLMMESHRRIVRCICPSRLTYLR
OAKVLCOLDEEVLHSPRLTNSAMRAGHLLDLTKRGKGAIAFLSEIKENHPDVTY
VNGLDPDVFNPSGSMETSKTTECAAGISLOEINERKOKVILARCOOEHL
GLAETAEGLHQLEADHSHKREVSANHEVRLKDEMLSLHYSNMLQEKELASR
CSLSOELLYLKQELQKRAMVSSCELEQOGLKRADESDEBELNLRKENETLRS
LTFSLAEKDLLEQSLDEARSGOELVERHSIRERAVAEOROEYMEKEETLLQFO
KSMACOLYREKYNALQAOVCELOKEROAYARPSAOERISQSLVEKDSLRQVEEL
TPOVCELTQLOLAEPBGVKKOARVTEPCPEKOLRYRHAIQPRDNDSCIVSS
TBSQLSDLSATSSRETVDFSSSPAPPSQSLRYAVEDGEEPMFSSTCLIEPG
DPCALPGAKAGPHLDVLELDLPADLPOLSSQLPVSFGLDVSESGVLMRRPARIL
SOVTMLAFQDALLQISVIGNTLGIPIHRYTPGSADQMLRGTQIVMDYASR
PLEKALDETTLLEAVGLRVDGFCILSKVNTDGYKLLQDLKAVTSGDSFYR
VNLAMBRAGKELOVHCNEVLAHTDTPMFOGCGCMHAHRNASTMTDTAAGTIPYRSR
AQQQLIALIQDITQOCTVTRKSSSGPQCLVAVISMDKAKAPILSLSPRQDLDSRM
EGSSTCFMABSCITLVPTLVWPHRPARPVLVPRAGKILSEKLLQGFKCLLA
EYLSOBEYAMSGRDIIOGEVSGRCWVTHRAVSLMERKTHALLDVLDSVCTLH
RMDIPPIVHVSVMERAKKLLKQLOLGTSEELHLEARQEGDILRAPCLYSLAP
DQMSLDGLSLCVRQALADEQKKVWVTESPR"
273..533
/note="Region: CARD domain"
misc_feature
588..1433
/note="Region: coiled-coil domain"
misc_feature
1914..2165
/note="Region: PDZ domain"
misc_feature
2232..2438
/note="Region: SH3 domain"
misc_feature
2646..3203
/note="Region: GUK domain"
BASE COUNT 799 a 1187 c 1263 g 682 t
ORIGIN
Query Match 84.2%; Score 351; DB 9; Length 3931;
Best Local Similarity 100.0%; Pred. No. 8.8e-58;
Matches 351; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ATGGGGGAACCTGTGCGCCAGGGAAGTCCGCACTACGCGCACTGACGAGAGACACTGTGG 60
DB 207 ATGGGGGAACCTGTGCGCCAGGGAAGTCCGCACTACGCGCACTGACGAGAGACACTGTGG 266
QY 61 GAGATGATGAGAGCCACCGCCACAGAGATGTAAGTGTGCTGCGCCACCGCCCTCACC 120
DB 267 GAGATGATGAGAGCCACCGCCACAGAGATGTAAGTGTGCTGCGCCACCGCCCTCACC 326
QY 121 CCTTACTGCGCCAGGCAAGTGTGCTGCGCAGTGGACGAGAGAGAGAGTGTGCAAGC 180
DB 327 CCTTACTGCGCCAGGCAAGTGTGCTGCGCAGTGGACGAGAGAGAGTGTGCAAGC 386
QY 181 CCCCAGCTACCAACAGCGCCATGCGGCGCGGCACTTGTGATTTGCTGAAGACTCGA 240
DB 387 CCCCAGCTACCAACAGCGCCATGCGGCGCGGCACTTGTGATTTGCTGAAGACTCGA 446
QY 241 GGGAGAAACGGGGCAATCGCTTCTCTGAGAGAGCTGAAGTTCCAAACCTGACGCTTAC 300
DB 447 GGGAGAAACGGGGCAATCGCTTCTCTGAGAGAGCTGAAGTTCCAAACCTGACGCTTAC 506
QY 301 ACCCTGATCAACGGGGCTGACGCGATGATGACTTCACTTAAGCTTTAGCGGT 351
DB 507 ACCCTGATCAACGGGGCTGACGCGATGATGACTTCACTTAAGCGGT 557

```

```

REFERENCE
1 (bases 1 to 3772)
Mcallister-Lucas,L.M., Inohara,N., Lucas,P.C., Rutland,J.,
Benito,A., Li,Q., Chen,S., Chen,F.F., Yamaoka,S., Verma,I.M.,
Mak,T.W. and Nunez,G.
Bim1, a MAPK family member linking protein kinase C activation to
Bcl10-mediated NF-kappaB induction
J. Biol. Chem. 276 (33), 30589-30597 (2001)
21391892
PUBMED
11387339
2 (bases 1 to 3772)
Mcallister-Lucas,L.M., Lucas,P.C., Inohara,N., Chen,S., Chen,F.F.
and Nunez,G.
Direct Submission
Submitted (22-MAR-2001) Department of Pathology, University of
Michigan Medical School, 1500 E. Medical Center Dr., Ann Arbor, MI
48109, USA
FEATURES
source
Location/Qualifiers
1..3772
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
1..3772
/gene="Bim2"
304..3303
/gene="Bim2"
/codon_start=1
/product="Bim2"
/protein_id="AAK60137.1"
/db_xref="GI:14326097"
/translation="MAELCRMDSTLTALDEEMLMDMLSHRRCIVQICPSRLTYLR
OAKVLCOLDEEVLHSPRLTNSAMRAGHLLDLKARNGKGAIAFLSEIKENHPDVTY
VNGLDPDVFNPSGSMETSKTTECAAGISLOEINERKOKVILARCOOEHL
GLAETAEGLHQLEADHSHKREVSANHEVRLKDEMLSLHYSNMLQEKELASR
CSLSOELLYLKQELQKRAMVSSCELEQOGLKRADESDEBELNLRKENETLRS
LTFSLAEKDLLEQSLDEARSGOELVERHSIRERAVAEOROEYMEKEETLLQFO
KSMACOLYREKYNALQAOVCELOKEROAYARPSAOERISQSLVEKDSLRQVEEL
TPOVCELTQLOLAEPBGVKKOARVTEPCPEKOLRYRHAIQPRDNDSCIVSS
TBSQLSDLSATSSRETVDFSSSPAPPSQSLRYAVEDGEEPMFSSTCLIEPG
DPCALPGAKAGPHLDVLELDLPADLPOLSSQLPVSFGLDVSESGVLMRRPARIL
SOVTMLAFQDALLQISVIGNTLGIPIHRYTPGSADQMLRGTQIVMDYASR
PLEKALDETTLLEAVGLRVDGFCILSKVNTDGYKLLQDLKAVTSGDSFYR
VNLAMBRAGKELOVHCNEVLAHTDTPMFOGCGCMHAHRNASTMTDTAAGTIPYRSR
AQQQLIALIQDITQOCTVTRKSSSGPQCLVAVISMDKAKAPILSLSPRQDLDSRM
EGSSTCFMABSCITLVPTLVWPHRPARPVLVPRAGKILSEKLLQGFKCLLA
EYLSOBEYAMSGRDIIOGEVSGRCWVTHRAVSLMERKTHALLDVLDSVCTLH
RMDIPPIVHVSVMERAKKLLKQLOLGTSEELHLEARQEGDILRAPCLYSLAP
DQMSLDGLSLCVRQALADEQKKVWVTESPR"

```

```

QY 241 GGGAGAAACGGGGCAATCGCTTCTCTGAGAGAGCTTCAACACCTGACGCTTAC 300

```

Db	Accession	Source	Organism	Version	Keywords	Definition	Locust	Result
Db	544	GGGAAGATGAGGCAATTCCTTTCTGGAAAGCGCTGAAGTCCACCAACCGTGAGTCA	603					
Db	301	ACCGTGGTCAACCGGCTGACCGCTGATGTTGACTTCACTTAACCTTACGGGT	351					
Db	604	ACCGTGGTCACTGGGCTGCACTGACCTTGAACCTTCAAGCACTTCAAGCGGT	654					
RESULT 6	BC029102							
LOCUS	BC029102							
DEFINITION	BC029102							
ACCESSION	BC029102							
VERSION	BC029102.1							
KEYWORDS	MGc.							
SOURCE	Mus musculus (house mouse)							
ORGANISM	Mus musculus							
REFERENCE	Strausberg, R. L., Felsing, E. A., Grouse, L. H., Derge, J. G., Klausner, R. D., Collins, P. S., Wagner, L., Sherman, C. M., Schuler, G. D., Altschul, S. F., Zeeberg, B., Buetow, K. H., Schaefer, C. F., Bat, N. K., Hopkins, R. F., Jordan, H., Moore, T., Max, S. I., Wang, J., Hsieh, P., Datchenko, L., Marusina, K., Farmer, A. A., Rubin, G. M., Hong, L., Siepleton, M., Soares, M. B., Bonaldo, M. F., Casavant, T. L., Scherle, T. E., Brownstein, M. J., Uedlin, T. B., Toshitsuki, S., Carimni, P., Prange, C., Rana, S. S., Loquellano, N. A., Peters, G. J., Abramson, R. D., Mullaly, S. J., Bosak, S. A., Mobwan, P. J., McKernan, R. J., Malek, J. A., Gunaratne, P. H., Richards, S., Willeay, K. C., Hale, S., Garcia, A. M., Gay, L. J., Hulyk, S. W., Villalón, D. K., Muzny, D. M., Sodergren, E. J., Lu, X., Gibbs, R. A., Pauley, J., Helton, E., Kettelman, M., Madan, A., Rodriguez, S., Sanchez, A., Whiting, M., Madan, A., Young, A. C., Shevchenko, Y., Bonfield, G. G., Blakesley, R. W., Touchman, J. W., Green, E. D., Dickson, M. C., Rodriguez, A. C., Grimwood, J., Schmutz, R., Myers, R. M., Butterfield, Y. S., Krzyzanski, M. I., Skalka, U., Smolins, D. E., Scherch, A., Schein, J. E., Jones, S. J., and Marra, M. A.							
TITLE	Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences							
JOURNAL	Proc. Natl. Acad. Sci. U.S.A.							
MEDLINE	12477932							
PUBMED	22388257							
REFERENCE	2 (bases 1 to 3995)							
AUTHORS	Strausberg, R.							
TITLE	Direct Submission							
JOURNAL	Submitted (01-MAY-2002) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA							
REMARK	NIH-MGC Project URL: <a href="http://mgc.ncl.nih.gov">http://mgc.ncl.nih.gov</a>							
COMMENT	Contact: MGC help desk Email: <a href="mailto:cgabs-remail.nih.gov">cgabs-remail.nih.gov</a> Tissue Procurement: Gilbert Smith, Ph.D. cDNA Library Preparation: Life Technologies, Inc. cDNA Library Arrayed by: The I.M.A.G.E Consortium (LLNL) DNA Sequencing by: Baylor College of Medicine Human Genome Sequencing Center Center code: BCM-HGSC Web site: <a href="http://www.hgsc.bcm.tmc.edu/cdna/">http://www.hgsc.bcm.tmc.edu/cdna/</a> Contact: <a href="mailto:amg@bcm.tmc.edu">amg@bcm.tmc.edu</a> Gunaratne, P. H., Garcia, A. M., Lu, X., Hulyk, S. W., Lounsberg, H., Kowis, C. R., Sneed, A. J., Martin, R. G., Muzny, D. M., Nanavati, A. N., Gibbs, R. A.							
FEATURES	Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E Consortium/LLNL at: <a href="http://image.llnl.gov">http://image.llnl.gov</a> Series: IRAX Plate: 36 Row: e Column: 1 This clone was selected for full length sequencing because it passed the following selection criteria: Similarity but not identity to protein. Location/Qualifiers 1..3995							

/organism="Mus musculus"  
 /mol\_type="mRNA"  
 /strain="FVB/N"  
 /db\_xref="taxon:10090"  
 /clone="MGC:28122 IMAGE:3979863"  
 /tissue type="Mammary tumor. Metallothionien-TGF alpha  
 model. 10 month old virgin mouse. Taken by biopsy."  
 /clone\_1ib="NCI\_CGAP\_Mam1"  
 /lab\_host="DH10B"  
 /note="Vector: pCMV-Sport6"  
 1. 3995  
 /gene="Card14"  
 /note="Synonyms: Bimp2, CARMA2"  
 /db\_xref="Locusid:170720"  
 /db\_xref="GI:22137688"  
 /db\_xref="MGI:2386258"  
 325\_..3324  
 /codon\_start=1  
 /product="caspase recruitment domain family, member 14"  
 /protein\_id="AAH29102.1"  
 /db\_xref="GI:22137688"  
 /db\_xref="LocusID:170720"  
 /translation="MAELIKMSDSTLRPLDEPMMDMLSHRCRIYVSCISRLPLPIR  
 QAVYLQDDEEILHSSRFNFSARVGHLLDLKARKKNAIAFLESLEKHNPDVYLL  
 VTGQSDIDFSTSGLMETSKEITCECLAISLSQELAEKQAEVLLRCCQKELR  
 GLAEHAEGKROLEVDHSMKREKSTGFHEVLKIDEMNLISYSALEKEELAIR  
 CHSLOEHLVKROELORASIVSQCERSRERSLKMSNLSPQGEELRLKEBNKILRS  
 MTSLEKXDLBESLDEBARSKEPLDRHISLSLRVAARQKQVTEHEKQTLQRR  
 KITVDCELYEKRMIMLQGVAELOKEDQVYARDAQMETISQRLVKEHLRRVPEL  
 TEQVLEKTLQRLIOAEAPGPKQOEAQELCRGQRLVRMAVCPDSDSCILSS  
 TBSRLMDLNTSREOMDFRSPSPMPPOQSLYKRVADFLDPESLSEPEVLEMR  
 LQAVYDPTDLEPEMIDGALSCQEDSLQSSRLSNVSESVPRVRRPARKTISQV  
 TVLAFQDALLBQIVYIGNLTGFIPIHRTVPGSAADMLRPGQIMNDYKPKPSEL  
 RATLENTLEQAVGLLRVNGSCYLSKLTITGEKXNLIQDLDAVNVTSQS FYIRVLD  
 AMORGDEGLQTHNDPLHTIDTFQORSQMAHNAVMPYTMKEDEPTIRYSAQOOQ  
 ILALIOMTQRCVPRKPPGSGPOKLVIVSVDRKVAPLTSSFDQSDMSKEEGGSS  
 VCMRSSECTPLAITYLLVHPHRRPARPRVLPVPLVGRILKKLCILQFPOCSAEYLS  
 QEYRATVSQKGDILIOGESIGIDHMTIRHAYVESIMMNSTALLDVRDSTRVLRMDM  
 PFIITHSVNKTAKRKRLKRLHRLGSSSEDFLEVARQEBLDRVPCLYSLADPSMS  
 DLSLSCTVLAIALDEKXKVTWTEPSC"  
 BASE COUNT 977 a 1095 c 1160 g 763 t  
 ORIGIN  
 Query Match 63.5%; Score 264.6; DB 10; Length 3995;  
 Best Local Similarity 84.6%; Pred. No. 3.6e-41;  
 Matches 297; Conservative 0; Mismatches 54; Indels 0; Gaps 0;  
 Db 325 ATGGCAAACTGTGTCGCCAGGAGACTCCGCACTACGGCACTGACGAGAGACACTGTGG 60  
 QY 1 ATGGGGAACTGTGTCGCCAGGAGACTCCGCACTACGGCACTGACGAGAGACACTGTGG 60  
 Db 325 ATGGCAAACTGTGTCGCCAGGAGACTCCGCACTGCACTGCTCTGGAGCAGAGATCTCTGG 384  
 QY 61 GAGATGATGAGAGACCCACCGCCACAGAGATGTGACGTGCAATGCCCCAGCGGCTCAAC 120  
 Db 385 GATATGCTGGAGATGACCCAGATGAGAGATTGTGAGAGACATGTGCCCTAGCGGCTCAAC 444  
 QY 121 CCTCACTGCGCCAGGCGCAAGTGTCTGTGCAGCTGACGAGAGAGAGAGTGTCTCAACAGC 180  
 Db 445 CCTCACTGCGCCAGGCGCAAGTGTCTGTGCAGCTGACGAGAGAGAGATCTCTCAATAGC 504  
 QY 181 CCCCAGCTCAACAACAAGCGGCATCGCGGACCGGCGGCACTTGCTGATTTGCTGAAGACTGA 240  
 Db 505 TCCGTTTCAACAACAATGCGCATAGAGATTGGCAGCTTGCTGATTTGCTGAAGAGCCGGA 564  
 QY 241 GGAAGAAGACGGGCGCATCGCTTCTGTGAAGAGCTGAAGTTCCAAACCTGACGTCTAC 300  
 Db 565 GGAAGAAGATGAGACCATGCTTCTGTGAAGAGCTGAAGTTCCAAACCTGACGTCTAC 624  
 QY 301 ACCCTGATCAACGGGCTGACGCTGAGTTGACTTCACTTAAGTAACTTAAGCGGT 351  
 Db 625 ACCCTGATCAACGGGCTGACGCTGAGTTGACTTCACTTAAGTAACTTAAGCGGT 675

LOCUS AC132816 55173 bp DNA linear HTG 04-SEP-2002  
DEFINITION Homo sapiens chromosome 17 clone RP13-593A24 map 17, LOW-PASS  
SEQUENCE SAMPLING.  
AC132816  
AC132816  
VERSION AC132816.1 GI:22711605  
KEYWORDS HTG; HTGS PHASEO.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE 1 (bases 1 to 55173)  
Barren, B., Nusbaum, C. and Lander, E.  
Homo sapiens chromosome 17, clone RP13-593A24  
Unpublished  
2 (bases 1 to 55173)  
Barren, B., Bastien, V., Bloom, T., Boguslavsky, L., Boukhgalter, B.,  
Barna, N., Bastien, V., Bloom, T., Choquet, Y., Collymore, A.,  
Camata, J., Chang, J., Chazaro, B., Dewar, K., Diaz, J.S., Dodge, S.,  
Cook, A., Cooke, P., Dearrellano, K., Dewar, K., Diaz, J.S., Dodge, S.,  
Faro, S., Ferreira, P., Fitzgerald, M., Gage, D., Galagan, J.,  
Gardyna, S., Gord, S., Graham, L., Grand-Pierre, N., Hagos, B.,  
Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C., Kamat, A.,  
Karatas, A., Kells, C., Landers, T., Levine, R., Lindblad-Toh, K.,  
Liu, G., Maclean, C., Macdonald, P., Major, J., Matthews, C.,  
McCarthy, M., Meldrum, J., Menes, L., Mihova, T., Mlenga, V.,  
Murphy, T., Naylor, J., Nguyen, C., Nicol, R., Norbu, C., Norman, C.H.,  
O'Connor, T., O'Donnell, P., O'Neill, D., Oliver, J., Peterson, K.,  
Plunkhang, P., Pierre, N., Raymond, C., Retta, R., Rise, C., Rogov, P.,  
Roman, J., Roy, A., Schauer, S., Schupbach, R., Seaman, S., Severy, P.,  
Smith, C., Spencer, B., Stange-Thomann, N., Stojanovic, N., Talamas, J.,  
Teifayez, S., Theodore, J., Topham, K., Travers, M., Vassiliev, H.,  
Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Young, G., Zaitoun, J.,  
Zemke, L., Zimmer, A. and Zody, M.  
Direct Submission  
Submitted (04-SEP-2002) Whitehead Institute/MIT Center for Genome  
Research, 320 Charles Street, Cambridge, MA 02141, USA  
All repeats were identified using RepeatMasker:  
Smit, A.P.A. & Green, P. (1996-1997)  
<http://ftp.genome.washington.edu/RM/RepeatMasker.html>  
----- Genome Center  
Center: Whitehead Institute/ MIT Center for Genome Research  
Center code: WIBR  
Web site: <http://www-seq.wi.mit.edu>  
Contact: [sequence.submissions@genome.wi.mit.edu](mailto:sequence.submissions@genome.wi.mit.edu)  
----- Project Information  
Center project name: L28179  
Center clone name: 593\_A\_24  
\* NOTE: This record contains 70 individual  
\* sequencing reads that have not been assembled into  
\* contigs. Runs of N are used to separate the reads  
\* and the order in which they appear is completely  
\* arbitrary. Low-pass sequence sampling is useful for  
\* identifying clones that may be gene-rich and allows  
\* overlap relationships among clones to be deduced.  
\* However, it should not be assumed that this clone  
\* will be sequenced to completion. In the event that  
\* the record is updated, the accession number will  
\* be preserved.  
\* 1 692: contig of 692 bp in length  
\* 693 792: gap of 100 bp  
\* 793 1479: contig of 687 bp in length  
\* 1480 1579: gap of 100 bp  
\* 1580 2271: contig of 692 bp in length  
\* 2272 2371: gap of 100 bp  
\* 2372 3056: contig of 685 bp in length  
\* 3057 3156: gap of 100 bp  
\* 3157 3858: contig of 702 bp in length  
\* 3859 3958: gap of 100 bp  
\* 3959 4632: contig of 674 bp in length  
\* 4633 4733: gap of 100 bp  
\* 4733 5418: contig of 686 bp in length  
\* 5419 5518: gap of 100 bp

5519 6199: contig of 681 bp in length  
6200 6299: gap of 100 bp in length  
6300 6969: contig of 670 bp in length  
6970 7069: gap of 100 bp in length  
7070 7779: contig of 710 bp in length  
7780 7879: gap of 100 bp in length  
7880 8569: contig of 689 bp in length  
8569 8669: gap of 100 bp in length  
8669 9364: contig of 636 bp in length  
9365 9464: gap of 100 bp in length  
9465 10171: contig of 707 bp in length  
10172 10271: gap of 100 bp in length  
10272 10963: contig of 692 bp in length  
10964 11063: gap of 100 bp in length  
11064 11751: contig of 688 bp in length  
11752 11851: gap of 100 bp in length  
11852 12547: contig of 696 bp in length  
12548 12647: gap of 100 bp in length  
12648 13309: contig of 662 bp in length  
13310 13409: gap of 100 bp in length  
13410 14098: contig of 689 bp in length  
14099 14198: gap of 100 bp in length  
14199 14869: contig of 671 bp in length  
14870 14969: gap of 100 bp in length  
14970 15667: contig of 698 bp in length  
15668 15767: gap of 100 bp in length  
15768 16479: contig of 712 bp in length  
16480 16579: gap of 100 bp in length  
16580 17268: contig of 689 bp in length  
17269 17368: gap of 100 bp in length  
17369 18038: contig of 670 bp in length  
18039 18138: gap of 100 bp in length  
18139 18829: contig of 691 bp in length  
18830 18929: gap of 100 bp in length  
18930 19632: contig of 703 bp in length  
19633 19732: gap of 100 bp in length  
19733 20432: contig of 700 bp in length  
20433 20532: gap of 100 bp in length  
20533 21213: contig of 681 bp in length  
21214 21313: gap of 100 bp in length  
21314 21988: contig of 675 bp in length  
21989 22088: gap of 100 bp in length  
22089 22801: contig of 713 bp in length  
22802 22901: gap of 100 bp in length  
22902 23575: contig of 674 bp in length  
23576 23575: gap of 100 bp in length  
23576 24355: contig of 680 bp in length  
24356 24455: gap of 100 bp in length  
24456 25146: contig of 691 bp in length  
25147 25246: gap of 100 bp in length  
25247 25947: contig of 701 bp in length  
25948 26047: gap of 100 bp in length  
26048 26740: contig of 693 bp in length  
26741 26840: gap of 100 bp in length  
26841 27543: contig of 703 bp in length  
27544 27643: gap of 100 bp in length  
27644 28338: contig of 695 bp in length  
28339 28438: gap of 100 bp in length  
28439 29137: contig of 699 bp in length  
29138 29237: gap of 100 bp in length  
29239 29924: contig of 687 bp in length  
29930 30024: gap of 100 bp in length  
30025 30712: contig of 688 bp in length  
30713 30812: gap of 100 bp in length  
30813 31504: contig of 692 bp in length  
31505 31604: gap of 100 bp in length  
31605 32322: contig of 718 bp in length  
32323 32422: gap of 100 bp in length  
32423 33104: contig of 682 bp in length  
33105 33204: gap of 100 bp in length  
33205 33899: contig of 695 bp in length  
33900 34690: gap of 100 bp in length  
34690 34690: contig of 691 bp in length

```

* 34691 34790: gap of 100 bp
* 34791 35481: contig of 691 bp in length
* 35481 35581: gap of 100 bp
* 35581 36277: contig of 696 bp in length
* 36277 36377: gap of 100 bp
* 36377 37072: contig of 695 bp in length
* 37072 37172: gap of 100 bp
* 37172 37856: contig of 684 bp in length
* 37856 37957: gap of 100 bp
* 37957 38652: contig of 696 bp in length
* 38652 38752: gap of 100 bp
* 38752 39440: contig of 688 bp in length
* 39440 39541: gap of 100 bp
* 39541 40218: contig of 678 bp in length
* 40218 40318: gap of 100 bp
* 40318 41015: contig of 697 bp in length
* 41015 41115: gap of 100 bp
* 41115 41820: contig of 705 bp in length
* 41820 41920: gap of 100 bp
* 41920 42600: contig of 680 bp in length
* 42600 42700: gap of 100 bp
* 42700 43386: contig of 686 bp in length
* 43386 43487: gap of 100 bp
* 43487 44179: contig of 693 bp in length
* 44179 44279: gap of 100 bp
* 44279 44939: contig of 660 bp in length
* 44939 45039: gap of 100 bp
* 45039 45703: contig of 664 bp in length
* 45703 45803: gap of 100 bp
* 45803 46470: contig of 667 bp in length
* 46470 46570: gap of 100 bp
* 46570 47264: contig of 694 bp in length
* 47264 47365: gap of 100 bp
* 47365 48069: contig of 705 bp in length
* 48069 48169: gap of 100 bp
* 48169 48847: contig of 678 bp in length
* 48847 48947: gap of 100 bp
* 48947 49636: contig of 689 bp in length
* 49636 49736: gap of 100 bp
* 49736 50437: contig of 701 bp in length
* 50437 50537: gap of 100 bp
* 50537 51213: contig of 676 bp in length
* 51213 51313: gap of 100 bp
* 51313 51999: contig of 686 bp in length
* 51999 52009: gap of 100 bp
* 52009 52792: contig of 693 bp in length
* 52792 52892: gap of 100 bp
* 52892 53594: contig of 702 bp in length
* 53594 53694: gap of 100 bp
* 53694 54377: contig of 683 bp in length
* 54377 54477: gap of 100 bp
* 54477 55173: contig of 696 bp in length.

```

FEATURES  
source 1. 55173

Query Match 50.8%; Score 212; DB 2; Length 55173;  
Best Local Similarity 100.0%; Pred. No. 3.5e-31;  
Matches 212; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 1 ATGGGGGAACTGTCGCCAGGAGCTCCGCACTCAAGCACTGACAGAGAGACAATCTGTG 60
DB 51604 ATGGGGGAACTGTCGCCAGGAGCTCCGCACTCAAGCACTGACAGAGAGACAATCTGTG 51663
QY 61 GAGATGATGAGAGACCCGCAAGAGATGATGATGATGATGATGATGATGATGATGATGATGATG 120
DB 51664 GAGATGATGAGAGACCCGCAAGAGATGATGATGATGATGATGATGATGATGATGATGATGATG 51723
QY 121 CCGTACTGCGCCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 180
DB 51724 CCGTACTGCGCCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 51783
QY 181 CCGCGGCTCACCAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 212

```

DB 51784 CCGCGGCTCACCAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 51815

RESULT 8  
AC123764  
LOCUS  
DEFINITION  
AC123764  
ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM  
Homo sapiens (human)  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
REFERENCE  
AUTHORS  
1 (bases 1 to 144000)  
Birren, B., Nusbaum, C., and Lander, E.  
Homo sapiens chromosome 17, clone CTC-784N20  
Unpublished  
2 (bases 1 to 144000)

Birren, B., Linton, L., Nusbaum, C., Lander, E., Ali, A., Allen, N.,  
Anderson, S., Bana, N., Bastien, V., Bloom, T., Boguslavsky, L.,  
Bouhgalter, B., Brown, A., Camarata, J., Campiano, A., Chang, J.,  
Chazaro, B., Choepel, Y., Colangelo, M., Collins, S., Collymore, A.,  
Cook, A., Cooke, P., Dearellano, K., Dewar, K., Diaz, J., Dodge, S.,  
Faro, S., Ferreira, P., Fitzgerald, M., Fitzhugh, W., Gage, D.,  
Galagan, J., Gardyna, S., Ginde, S., Gord, S., Goyette, M., Graham, L.,  
Grand-Pierre, N., Hagos, B., Horton, L., Hulme, W., Iliev, I.,  
Johnson, R., Jones, C., Kamat, A., Karatas, A., Kells, C., Lacroque, K.,  
Lamazares, R., Landers, T., Lehoczy, J., Levine, R., Lindblad-Toh, K.,  
Liu, G., Maclean, C., Macdonald, P., Major, J., Marcus, N.,  
Matthews, C., McCarthy, M., McSwan, P., McKernan, K., Meldrum, J.,  
Menus, L., Minova, T., Mienga, V., Murphy, T., Naylor, J., Nguyen, C.,  
Nicol, R., Noibu, C., Norman, C. H., O'Connor, T., O'Donnell, P.,  
O'Neill, D., Oliver, J., Peterson, K., Punthang, P., Pierre, N.,  
Pollara, P., Raymond, C., Retta, R., Rieback, M., Riley, R., Rise, C.,  
Rogov, P., Roman, J., Rosetti, M., Roy, A., Santos, R., Schauer, S.,  
Schupack, R., Seaman, S., Severy, P., Spencer, B., Stange-Thomann, N.,  
Stojanovic, N., Strauss, N., Subramanian, A., Talamas, J., Tesfaye, S.,  
Theodore, J., Topham, K., Travers, N., Travis, N., Triggillo, J.,  
Vassiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W., J.,  
Young, G., Zainoun, J., Zemdek, L., Zimmer, A. and Zody, M.

Direct Submission  
Submitted (01-JUN-2002) Whitehead Institute/MIT Center for Genome  
Research, 320 Charles Street, Cambridge, MA 02141, USA  
3 (bases 1 to 144000)

REFERENCE  
AUTHORS  
Birren, B., Nusbaum, C., Lander, E., Ali, A., Allen, N., Anderson, S.,  
Bana, N., Bastien, V., Bloom, T., Boguslavsky, L., Bouhgalter, B.,  
Camarata, J., Chang, J., Chazaro, B., Choepel, Y., Collymore, A.,  
Cook, A., Cooke, P., Dearellano, K., Dewar, K., Diaz, J., Dodge, S.,  
Faro, S., Ferreira, P., Fitzgerald, M., Gage, D., Galagan, J.,  
Gardyna, S., Gord, S., Graham, L., Grand-Pierre, N., Hafez, N.,  
Hagos, B., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C.,  
Kamat, A., Karatas, A., Kells, C., Landers, T., Levine, R.,  
Lindblad-Toh, K., Liu, G., Maclean, C., Macdonald, P., Major, J.,  
Matthews, C., McCarthy, M., Meldrum, J., Menus, L., Minova, T.,  
Mienga, V., Murphy, T., Naylor, J., Nguyen, C., Nicol, R., Noibu, C.,  
Norman, C. H., O'Connor, T., O'Donnell, P., O'Neill, D., Oliver, J.,  
Peterson, K., Punthang, P., Pierre, N., Raymond, C., Retta, R.,  
Rise, C., Rogov, P., Roman, J., Roy, A., Schauer, S., Schupack, R.,  
Seaman, S., Severy, P., Smith, C., Spencer, B., Stange-Thomann, N.,  
Stojanovic, N., Talamas, J., Tesfaye, S., Theodore, J., Topham, K.,  
Travers, N., Vassiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X.,  
Wyman, D., Young, G., Zainoun, J., Zemdek, L., Zimmer, A. and Zody, M.

Direct Submission  
Submitted (16-JAN-2003) Whitehead Institute/MIT Center for Genome  
Research, 320 Charles Street, Cambridge, MA 02141, USA  
4 (bases 1 to 144000)

REFERENCE  
AUTHORS  
Birren, B., Nusbaum, C., Lander, E., Abouelelli, A., Allen, N.,  
Anderson, S., Arachchi, H. M., Bana, N., Bastien, V., Bloom, T.,  
Boguslavsky, L., Bouhgalter, B., Camarata, J., Chang, J., Choepel, Y.,  
Collymore, A., Cook, A., Cooke, P., Corum, B., Dearellano, K.,  
Diaz, J., Dodge, S., Dooley, K., Dorris, L., Erickson, J., Faro, S.,  
Ferreira, P., Fitzgerald, M., Gage, D., Galagan, J., Gardyna, S.,

Graham, L., Grand-Pierre, N., Hafez, N., Hagopian, D., Hagos, B., Hall, J., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C., Kamat, A., Karatas, A., Kells, C., Landers, T., Levine, R., Lindblad-Toh, K., Liu, G., Lui, A., Mabbitt, R., Maclean, C., Macdonald, P., Major, J., Manning, J., Matthews, C., McCarthy, M., Meldrim, J., Menus, L., Mihova, T., Mlenga, A., Murphy, T., Naylor, J., Nguyen, C., Nicol, R., Norbu, C., O'Connor, T., O'Donnell, P., O'Neill, D., Oliver, J., Peterson, K., Phunkhang, P., Pierre, N., Rachupka, A., Ramasamy, U., Raymond, C., Retta, R., Rise, C., Rogov, P., Roman, J., Schauer, S., Schupack, R., Seaman, S., Severy, P., Smith, C., Spencer, B., Stange-Thomann, N., Stojanovic, N., Stubbs, M., Talamas, J., Testaye, S., Theodore, J., Topham, K., Travers, M., Vassiliev, H., Venkataraman, V., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.

Direct Submission  
Submitted (26-JAN-2003) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA  
5 (bases 1 to 144000)

REFERENCE  
AUTHORS

TITLE  
JOURNAL  
COMMENT

Submitted (11-MAR-2003) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA  
On Mar 11, 2003 this sequence version replaced gi:27902327.  
All repeats were identified using RepeatMasker:  
Smit, A.F.A. & Green, P. (1996-1997)  
<http://ftp.genome.washington.edu/RM/RepeatMasker.html>

Center: Whitehead Institute/MIT Center for Genome Research  
Center code: WIRB  
Web site: <http://www-seq.wi.mit.edu>  
Contact: [sequence\\_submissions@genome.wi.mit.edu](mailto:sequence_submissions@genome.wi.mit.edu)  
----- Project Information  
Center project name: L27277  
Center clone name: 784\_N\_N\_20

Only the first 144,0 kilobases of this clone are being submitted.  
The remainder overlaps accession number AC124319 [WIGR project L27275].

FEATURES  
source

1. 144000  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"  
/chromosome="17"  
/map="17"  
/clone="CTC-784N20"  
/clone\_lib="CITC Human BAC"  
212. 764  
/rpt\_family="MIR"  
repeat\_region  
complement(381..1030)  
/rpt\_family="L2"  
1977..2009  
/rpt\_family="(TG)n"

	repeat_region	2196..2304
	repeat_region	/rpt_family="(TCCCC)n"
	repeat_region	complement(2305..2587)
	repeat_region	/rpt_family="AluX"
	repeat_region	2618..2752
	repeat_region	/rpt_family="CT-rich"
	repeat_region	complement(2756..2929)
	repeat_region	/rpt_family="FRAM"
	repeat_region	2939..2962
	repeat_region	/rpt_family="(TTTGG)n"
	repeat_region	complement(4091..4184)
	repeat_region	/rpt_family="MER94"
	repeat_region	complement(4330..4638)
	repeat_region	/rpt_family="AluX"
	repeat_region	4708..4756
	repeat_region	/rpt_family="CT-rich"
	repeat_region	4760..4782
	repeat_region	/rpt_family="(CA)n"
	repeat_region	complement(4876..4946)
	repeat_region	/rpt_family="HAL1b"
	repeat_region	4976..5002
	repeat_region	/rpt_family="AT-rich"
	repeat_region	complement(5003..5050)
	repeat_region	/rpt_family="MER68A"
	repeat_region	complement(5058..5349)
	repeat_region	/rpt_family="AluSg"
	repeat_region	5350..5382
	repeat_region	/rpt_family="AT-rich"
	repeat_region	5510..5651
	repeat_region	/rpt_family="BC200"
	repeat_region	complement(5658..5963)
	repeat_region	/rpt_family="AluX"
	repeat_region	complement(6417..6730)
	repeat_region	/rpt_family="AluDb"
	repeat_region	complement(7421..7543)
	repeat_region	/rpt_family="FLAM-C"
	repeat_region	complement(7560..7591)
	repeat_region	/rpt_family="L1MB8"
	repeat_region	7592..7909
	repeat_region	/rpt_family="AluDo"
	repeat_region	complement(7910..8041)
	repeat_region	/rpt_family="L1MB8"
	repeat_region	8112..8231
	repeat_region	/rpt_family="MER45A"
	repeat_region	complement(8406..8661)
	repeat_region	/rpt_family="MLT1C"
	repeat_region	complement(8662..8856)
	repeat_region	/rpt_family="AluDb"
	repeat_region	8857..9148
	repeat_region	/rpt_family="AluX"
	repeat_region	complement(9149..9286)
	repeat_region	/rpt_family="AluDb"
	repeat_region	complement(9287..9441)
	repeat_region	/rpt_family="MLT1C"
	repeat_region	complement(10183..10488)
	repeat_region	/rpt_family="AluX"
	repeat_region	10638..10845
	repeat_region	/rpt_family="L1MB3B"
	repeat_region	11016..11282

Query Match 50.8%; Score 212; DB 9; Length 144000;  
Best Local Similarity 100.0%; Pred. No. 3.2e-31;  
Matches 212; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	ATGGGGGAACCTGCGCGAGGAGCTCCGACATCAAGGACATGAGAGAGCACTGTGG	60
DB	12882	ATGGGGGAACCTGCGCGAGGAGCTCCGACATCAAGGACATGAGAGAGCACTGTGG	12941
QY	61	GAGATGATGAGAGCCACCGCCACGATCGTACGCTGATCTGCCACCGCCTCAAC	120
DB	12942	GAGATGATGAGAGCCACCGCCACGATCGTACGCTGATCTGCCACCGCCTCAAC	13001
QY	121	CCCTACCTGCGCGAGGAGAGGTGCTGCGACGCTGAGAGAGAGAGAGTGTGACAGC	180







```

repeat_region complement(10268. 10576)
/rpt_family="AluXs"
repeat_region complement(10582. 10738)
/rpt_family="FRAM"
repeat_region complement(10806. 10962)
/rpt_family="AluSg"
unsure complement(10958. 10973)
/note="single clone coverage"
repeat_region 11049. 11236
/rpt_family="AluDb"
repeat_region 12381. 12509
/rpt_family="Charlie8"
repeat_region complement(12611. 12920)
/rpt_family="AluSg"
repeat_region complement(12949. 13219)
/rpt_family="AluSc"
repeat_region complement(15127. 15161)
/rpt_family="L2"
repeat_region complement(15128. 15283)
/rpt_family="MIR"
repeat_region 15284. 15647
/rpt_family="HELIC"
repeat_region complement(15648. 15685)
/rpt_family="MIR"
repeat_region complement(15782. 15860)
/rpt_family="L2"
repeat_region 17124. 17422
/rpt_family="AluXs"
repeat_region complement(17967. 18101)
/rpt_family="L1MC4A"
repeat_region complement(19532. 19840)
/rpt_family="AluXs"
repeat_region complement(21106. 21402)
/rpt_family="AluDb"
repeat_region 22639. 23145
/rpt_family="L1MC4a"
repeat_region 23146. 23203
/rpt_family="(TATATG)n"
repeat_region 23204. 23323
/rpt_family="L1MC4a"
repeat_region 23324. 23637
/rpt_family="AluXs"
repeat_region 23638. 23851
/rpt_family="L1MC4a"
repeat_region 24094. 24625
/rpt_family="L1MC4a"
repeat_region 24826. 25117
/rpt_family="AluDb"
repeat_region complement(25762. 25901)
/rpt_family="L2"
repeat_region 26264. 26442
/rpt_family="(TGGG)n"
repeat_region complement(28083. 28218)
/rpt_family="MIR"
repeat_region 28255. 28430
/rpt_family="MLTID"
repeat_region 28783. 28913
/rpt_family="HERA1D"
repeat_region 29046. 29084
/rpt_family="AT_rich"
repeat_region 29111. 29392
/rpt_family="AluDb"
repeat_region 29396. 29433

```

1 ATGGGGGAACTGTGCGCGAGGAGCTCGCAGCTACCGGCAGTGGACGAGGAGACACTGTGG 60  
 28734 ATGGGGAACTGTGCGCGAGGAGCTCGCAGCTACCGGCAGTGGACGAGGAGACACTGTGG 28675

QY	Db	QY	Db	RESULT 10
61	GAGATATGAGAGCCACCGCCACAGAGATCGTATCGTATGTCGCCCGCCCTCACC	181	CCCCGCTCACCAACAGCGCCATGCGAGCCCGG	AC015559
28674	GAGATATGAGAGCCACCGCCACAGAGATCGTATGTCGCCCGCCCTCACC	28554	CCCCGCTCACCAACAGCGCCATGCGAGCCCGG	AC015559
121	CCCTACCTGTGGCCAGGCGCAAGTGTCTGTGCCAGTGTGACGAGAGAGAGGTGTGACAGC			LOCUS
28614	CCTTACTGTGGCCAGGCGCAAGTGTCTGTGCCAGTGTGACGAGAGAGAGGTGTGACAGC			DEFINITION
				AC015559
				VERSION
				KEYWORDS
				SOURCE
				ORGANISM
				REFERENCE
				AUTHORS
				JOURNAL
				REFERENCE
				AUTHORS
				TITLE
				JOURNAL
				COMMENT

```

Center: Stanford DNA Sequencing and Technology Development
Center
Center code: SNSTC
Web site: http://sequence-www.stanford.edu/group/human/
Contact: hum-info@sequence.stanford.edu
-----
Project Information
Center project name: 708
Center clone name: RP11-334C17
-----
Summary Statistics
Sequencing Vector: M13mp18; X02513
Chemistry: Dye-Primer; 1% of reads
Chemistry: Dye-terminator Big Dye; 8% of reads
Assembly program: Phrap; version 0.990319
Consensus quality: 180520 bases at least Q40
Consensus quality: 181812 bases at least Q30
Consensus quality: 182303 bases at least Q20
Insert size: 190560; agarose-fp.
Insert size: 16865; sum-of-contigs
Quality coverage: 9.0x in Q20 bases; agarose-fp
Quality coverage: 9.2x in Q20 bases; sum-of-contigs.
NOTE: This is a "working draft" sequence. It currently
* consists of 11 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
      1
3078: contig of 3078 bp in length
3079      3178: gap of unknown length

```





*	11852	12547	contig of 666 bp	in length
*	12548	12647	gap of 100 bp	in length
*	12648	13309	contig of 662 bp	in length
*	13310	13409	gap of 100 bp	in length
*	13410	14098	contig of 689 bp	in length
*	14099	14198	gap of 100 bp	in length
*	14199	14689	contig of 671 bp	in length
*	14670	14699	gap of 100 bp	in length
*	14970	15667	contig of 698 bp	in length
*	15668	15767	gap of 100 bp	in length
*	15768	16479	contig of 712 bp	in length
*	16480	16579	gap of 100 bp	in length
*	16580	17268	contig of 689 bp	in length
*	17269	17368	gap of 100 bp	in length
*	17369	18038	contig of 670 bp	in length
*	18039	18138	gap of 100 bp	in length
*	18139	18829	contig of 691 bp	in length
*	18830	18929	gap of 100 bp	in length
*	18930	19632	contig of 703 bp	in length
*	19633	19732	gap of 100 bp	in length
*	19733	20432	contig of 700 bp	in length
*	20433	20532	gap of 100 bp	in length
*	20533	21213	contig of 681 bp	in length
*	21213	21213	gap of 100 bp	in length
*	21314	21988	contig of 675 bp	in length
*	21989	22088	gap of 100 bp	in length
*	22089	22801	contig of 713 bp	in length
*	22802	22901	gap of 100 bp	in length
*	22902	23575	contig of 674 bp	in length
*	23576	23675	gap of 100 bp	in length
*	23676	24355	contig of 680 bp	in length
*	24356	24455	gap of 100 bp	in length
*	24456	25146	contig of 691 bp	in length
*	25147	25246	gap of 100 bp	in length
*	25247	25947	contig of 701 bp	in length
*	25948	26047	gap of 100 bp	in length
*	26048	26740	contig of 693 bp	in length
*	26741	26840	gap of 100 bp	in length
*	26841	27543	contig of 703 bp	in length
*	27544	27643	gap of 100 bp	in length
*	27644	28338	contig of 695 bp	in length
*	28339	28438	gap of 100 bp	in length
*	28439	29137	contig of 699 bp	in length
*	29138	29237	gap of 100 bp	in length
*	29238	29924	contig of 687 bp	in length
*	29925	30024	gap of 100 bp	in length
*	30025	30812	contig of 688 bp	in length
*	30813	30912	gap of 100 bp	in length
*	30913	31504	contig of 692 bp	in length
*	31505	31604	gap of 100 bp	in length
*	31605	32322	contig of 718 bp	in length
*	32323	32423	gap of 100 bp	in length
*	32423	33104	contig of 682 bp	in length
*	33105	33204	gap of 100 bp	in length
*	33205	33899	contig of 695 bp	in length
*	33900	33999	gap of 100 bp	in length
*	34000	34690	contig of 691 bp	in length
*	34691	34790	gap of 100 bp	in length
*	34791	35481	contig of 691 bp	in length
*	35482	35581	gap of 100 bp	in length
*	35582	36272	contig of 696 bp	in length
*	36273	36377	gap of 100 bp	in length
*	36378	37072	contig of 695 bp	in length
*	37073	37172	gap of 100 bp	in length
*	37173	37856	contig of 684 bp	in length
*	37857	37956	gap of 100 bp	in length
*	37957	38652	contig of 696 bp	in length
*	38653	38752	gap of 100 bp	in length
*	38753	39440	contig of 688 bp	in length
*	39441	39540	gap of 100 bp	in length
*	39541	40212	contig of 678 bp	in length
*	40213	40312	gap of 100 bp	in length
*	40313	40512	contig of 657 bp	in length

FEATURES	Source	Location/Qualifiers
Query Match	49.2%; Score 205; DB 2; Length 55173;	
Best Local Similarity	100.0%; Pred. No. 7.9e-30;	
Matches 205; Conservative	0; Mismatches 0; Indels 0; Gaps 0;	
QY	211 GGGGCACTTGCTGGATTGCTGTAAGACTGAGGAGAAAGACGGGGCCATCGCCTTCTGGAG	270
DB	20167 GGGGCACTTGCTGGATTGCTGTAAGACTGAGGAGAAAGACGGGGCCATCGCCTTCTGGAG	20108
QY	271 AGGCTAAGTCCACAAACCCCTGACGCTCAACCCCTGGTACCGGGGTGAGGCTGATGTT	330
DB	20107 AGGCTAAGTCCACAAACCCCTGACGCTCAACCCCTGGTACCGGGGTGAGGCTGATGTT	20048
QY	331 GACTTGAAGTACTTTCAGCGGTGAGAGCTCCGACTTTGACGTTTGGCAGGCACTTCTAG	390
DB	20047 GACTTGAAGTACTTTCAGCGGTGAGAGCTCCGACTTTGACGTTTGGCAGGCACTTCTAG	19988
QY	391 AACCTGAGGCTCCTGTGAACCCGAG	415
DB	19987 AACCTGAGGCTCCTGTGAACCCGAG	19963
RESULT 13		
LOCUS	AC137736	
DEFINITION	Homo sapiens chromosome 17 clone RP13-365D24 map 17, LOW-PASS	
ACCESSION	AC137736	74301 bp
VERSION	AC137736.1	DNA
KEYWORDS	HTG; HTGS; PHAS80.	linear
SOURCE	Homo sapiens (human)	HTG 28-NOV-2002
ORGANISM	Homo sapiens	
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.	
	(bases 1 to 74301)	

AUTHORS Birren, B., Nusbaum, C. and Lander, E.  
 TITLE Homo sapiens chromosome 17, clone RP13-365D24  
 JOURNAL Unpublished  
 REFERENCE 2 (bases 1 to 74301)  
 AUTHORS Birren, B., Nusbaum, C., Lander, E., Ali, A., Allen, N., Anderson, S.,  
 Barra, N., Bastien, V., Bloom, T., Boguslavsky, L., Bokhgale, B.,  
 Camarata, J., Chang, J., Chazaro, B., Choepel, Y., Collymore, A.,  
 Cook, A., Cooke, P., DeArrelano, K., Dewar, K., Diaz, U. S., Dodge, S.,  
 Fairo, S., Ferrel, P., Fitzgerald, M., Gage, D., Galagan, J.,  
 Gaidyna, C., Goid, S., Graham, L., Grand-Pierre, N., Hafez, N.,  
 Hagos, B., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C.,  
 Kamet, A., Karatas, A., Kelle, C., Landers, T., Levine, R.,  
 Lindblad-Toh, K., Liu, G., Maclean, C., Macdonald, P., Major, J.,  
 Matthews, C., McCarthy, M., Meldrum, J., Meneus, L., Mihova, T.,  
 Mlenga, V., Murphy, T., Naylor, J., Nguyen, C., Nicol, R., Norbu, C.,  
 Norman, C. H., O'Connor, T., O'Donnell, P., O'Neill, D., Oliver, J.,  
 Peterson, K., Phunkhang, P., Pierre, N., Raymond, C., Retta, R.,  
 Rise, C., Rogov, P., Roman, J., Roy, A., Schauer, S., Schnupack, R.,  
 Seaman, S., Severy, P., Smith, C., Spencer, B., Stange-Thomann, N.,  
 Stojanovic, N., Talamas, J., Tesfaye, S., Theodore, J., Topham, K.,  
 Travers, M., Vassiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X.,  
 Wymann, D., Young, G., Zaitoun, J., Zembek, L., Zimmer, A. and Zody, M.

TITLE Direct Submission  
 JOURNAL Submitted (28-NOV-2002) Whitehead Institute/MIT Center for Genome  
 Research, 320 Charles Street, Cambridge, MA 02141, USA  
 COMMENT All repeats were identified using RepeatMasker:  
 Smit, A.F.A. & Green, P. (1996-1997)  
 http://ftp.genome.washington.edu/RM/RepeatMasker.html

Genome Center  
 Center: Whitehead Institute/ MIT Center for Genome Research  
 Center code: WIBR  
 Web site: http://www-seq.wi.mit.edu  
 Contact: sequence.submissions@genome.wi.mit.edu  
 Project Information  
 Center project name: L28728  
 Center clone name: 365\_D\_24

\* NOTE: This record contains 91 individual  
 \* sequencing reads that have not been assembled into  
 \* contigs. Runs of N are used to separate the reads  
 \* and the order in which they appear is completely  
 \* arbitrary. Low-pass sequence sampling is useful for  
 \* identifying clones that may be gene-rich and allows  
 \* overlap relationships among clones to be deduced.  
 \* However, it should not be assumed that this clone  
 \* will be sequenced to completion. In the event that  
 \* the record is updated, the accession number will  
 \* be preserved.

\* 1 707: contig of 707 bp in length  
 \* 708 807: gap of 100 bp  
 \* 808 1540: contig of 733 bp in length  
 \* 1541 1640: gap of 100 bp  
 \* 1641 2344: contig of 704 bp in length  
 \* 2345 2444: gap of 100 bp  
 \* 2445 3156: contig of 712 bp in length  
 \* 3157 3256: gap of 100 bp  
 \* 3257 3978: contig of 722 bp in length  
 \* 3979 4078: gap of 100 bp  
 \* 4079 4781: contig of 703 bp in length  
 \* 4782 4881: gap of 100 bp  
 \* 4882 5584: contig of 713 bp in length  
 \* 5595 5694: gap of 100 bp  
 \* 5695 6405: contig of 711 bp in length  
 \* 6406 6505: gap of 100 bp  
 \* 6506 7197: contig of 692 bp in length  
 \* 7198 7297: gap of 100 bp  
 \* 7298 8015: contig of 718 bp in length  
 \* 8016 8115: gap of 100 bp  
 \* 8116 8818: contig of 703 bp in length  
 \* 8819 8918: gap of 100 bp  
 \* 8919 9653: contig of 735 bp in length  
 \* 9654 9753: gap of 100 bp  
 \* 9754 10454: contig of 701 bp in length

\* 10455 10554: gap of 100 bp  
 \* 10555 11277: contig of 723 bp in length  
 \* 11278 11377: gap of 100 bp  
 \* 11378 12081: contig of 704 bp in length  
 \* 12082 12181: gap of 100 bp  
 \* 12182 12913: contig of 732 bp in length  
 \* 12914 13014: gap of 100 bp  
 \* 13015 13742: contig of 729 bp in length  
 \* 13743 13842: gap of 100 bp  
 \* 13843 14560: contig of 718 bp in length  
 \* 14561 14660: gap of 100 bp  
 \* 14661 15383: contig of 722 bp in length  
 \* 15384 15482: gap of 100 bp  
 \* 15483 16198: contig of 716 bp in length  
 \* 16199 16299: gap of 100 bp  
 \* 16299 17006: contig of 708 bp in length  
 \* 17007 17106: gap of 100 bp  
 \* 17107 17817: contig of 711 bp in length  
 \* 17818 17917: gap of 100 bp  
 \* 17918 18642: contig of 725 bp in length  
 \* 18643 18742: gap of 100 bp  
 \* 18743 19458: contig of 716 bp in length  
 \* 19459 19558: gap of 100 bp  
 \* 19559 20264: contig of 706 bp in length  
 \* 20265 21103: contig of 739 bp in length  
 \* 21104 21203: gap of 100 bp  
 \* 21204 21910: contig of 707 bp in length  
 \* 21911 22010: gap of 100 bp  
 \* 22011 22731: contig of 721 bp in length  
 \* 22732 22831: gap of 100 bp  
 \* 22832 23556: contig of 725 bp in length  
 \* 23557 23656: gap of 100 bp  
 \* 23657 24373: contig of 717 bp in length  
 \* 24374 24473: gap of 100 bp  
 \* 24474 25166: contig of 693 bp in length  
 \* 25167 25266: gap of 100 bp  
 \* 25267 25949: contig of 683 bp in length  
 \* 25950 26049: gap of 100 bp  
 \* 26050 26763: contig of 714 bp in length  
 \* 26764 26863: gap of 100 bp  
 \* 26864 27583: contig of 720 bp in length  
 \* 27584 28421: gap of 100 bp  
 \* 28422 28521: contig of 738 bp in length  
 \* 28522 29223: gap of 100 bp  
 \* 29224 29323: contig of 702 bp in length  
 \* 29324 30031: gap of 100 bp  
 \* 30032 30131: contig of 708 bp in length  
 \* 30132 30859: gap of 100 bp  
 \* 30860 30959: contig of 728 bp in length  
 \* 30960 31693: gap of 100 bp  
 \* 31694 31793: contig of 734 bp in length  
 \* 31794 32516: gap of 100 bp  
 \* 32517 32616: contig of 723 bp in length  
 \* 32617 33338: gap of 100 bp  
 \* 33339 33438: contig of 722 bp in length  
 \* 33439 34162: gap of 100 bp  
 \* 34163 34262: contig of 724 bp in length  
 \* 34263 34983: gap of 100 bp  
 \* 34984 35082: contig of 720 bp in length  
 \* 35083 35785: gap of 100 bp  
 \* 35786 35885: contig of 703 bp in length  
 \* 35886 36612: gap of 100 bp  
 \* 36613 36712: contig of 727 bp in length  
 \* 36713 37460: gap of 100 bp  
 \* 37461 37560: contig of 748 bp in length  
 \* 37561 38293: gap of 100 bp  
 \* 38294 38393: contig of 733 bp in length  
 \* 38394 39114: gap of 100 bp  
 \* 39115 39214: contig of 721 bp in length  
 \* 39215 39943: gap of 100 bp  
 \* 39944 40043: contig of 729 bp in length

```

* 40044 40763: contig of 720 bp in length
* 40764 40863: gap of 100 bp
* 40864 41562: contig of 699 bp in length
* 41563 41662: gap of 100 bp
* 41663 42386: contig of 724 bp in length
* 42387 42486: gap of 100 bp
* 42487 43202: contig of 716 bp in length
* 43203 43302: gap of 100 bp
* 43303 44007: contig of 705 bp in length
* 44008 44107: gap of 100 bp
* 44108 44794: contig of 687 bp in length
* 44795 44894: gap of 100 bp
* 44895 45581: contig of 687 bp in length
* 45582 45681: gap of 100 bp
* 45682 46363: contig of 682 bp in length
* 46364 46463: gap of 100 bp
* 46464 47166: contig of 703 bp in length
* 47167 47266: gap of 100 bp
* 47267 48002: contig of 736 bp in length
* 48003 48102: gap of 100 bp
* 48103 48831: contig of 729 bp in length
* 48832 48931: gap of 100 bp
* 48932 49651: contig of 720 bp in length
* 49652 49751: gap of 100 bp
* 49752 50488: contig of 737 bp in length
* 50489 50588: gap of 100 bp
* 50589 51295: contig of 707 bp in length
* 51296 51395: gap of 100 bp
* 51396 52123: contig of 728 bp in length
* 52124 52223: gap of 100 bp
* 52224 52935: contig of 712 bp in length
* 52936 53035: gap of 100 bp
* 53036 53743: contig of 708 bp in length
* 53744 53843: gap of 100 bp
* 53844 54539: contig of 695 bp in length
* 54539 54638: gap of 100 bp
* 54639 55354: contig of 716 bp in length
* 55355 55454: gap of 100 bp
* 55455 56159: contig of 705 bp in length
* 56160 56259: gap of 100 bp
* 56260 56986: contig of 727 bp in length
* 56987 57086: gap of 100 bp
* 57087 57807: contig of 721 bp in length

```

Query Match 45.3% Score 188.8; DB 2; Length 74301;  
 Best Local Similarity 62.3% Pred. No. 9.9e-27;  
 Matches 202; Conservative 0; Mismatches 122; Indels 0; Gaps 0;

```

QY 1 ATGGGAGAACTGTCCGAGGAGCTCCGACTCAGGCACTGAGAGAGACAATGTGG 60
DB 24185 ATGGGAGAACTGTCCGAGGAGCTCCGACTCAGGCACTGAGAGAGACAATGTGG 24244
QY 61 GAGATGATGAGAGACCAACCGCACAGAGATCGTACCTGATCCCGACCGCTTCAAC 120
DB 24245 GAGATGATGAGAGACCAACCGCACAGAGATCGTACCTGATCCCGACCGCTTCAAC 24304
QY 121 CCTTACTGCGCCGAGGAGAGAGTCTGTGCTGAGAGAGAGAGAGAGTCTGACACG 180
DB 24305 CCTTACTGCGCCGAGGAGAGAGTCTGTGCTGAGAGAGAGAGAGAGTCTGACACG 24364
QY 181 CCGCGGCTCACCACAGCGCATGCGGCGGAGCACTTGCTGATTCGAGAGACTCGA 240
DB 24365 CCGCGGCTCACCACAGCGCATGCGGCGGAGCACTTGCTGATTCGAGAGACTCGA 24424
QY 241 GGGAGAGACGGGAGCATGCGCTTCTGAGAGAGCTGAAGTTCCACAACCTGAGCTTAC 300
DB 24425 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 24484
QY 301 ACCCTGTGTACCGGCGCTGACAGCT 324
DB 24485 TATCTCGCCCTGTGTGATTCCT 24508

```

RESULT 14  
 AC132197/c  
 LOCUS  
 DEFINITION  
 ACCESSION  
 AC132197  
 VERSION  
 AC132197.1 GI:22549815  
 KEYWORDS  
 HTG; HTGS\_PHASE0.  
 SOURCE  
 ORGANISM  
 Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 1 (bases 1 to 65608)  
 Birren,B., Nusbaum,C., Lander,E., Ali,A., Allen,N., Anderson,S.,  
 Birren,B., Basteien,V., Bloom,T., Boguslavsky,L., Boukhgalter,B.,  
 Camarata,J., Chang,J., Chazaro,B., Choepel,Y., Collymore,A.,  
 Cook,A., Cooke,P., Dearellano,K., Dewar,K., Diaz,J.S., Dodge,S.,  
 Faro,S., Ferreira,P., Fitzgerald,M., Gage,D., Galagan,J.,  
 Gardyna,S., Gord,S., Graham,L., Grand-Pierre,N., Hagos,B.,  
 Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C., Kamat,A.,  
 Karatas,A., Kelle,C., Landers,T., Levine,R., Lindblad-Toh,K.,  
 Liu,G., Maclean,C., Macdonald,P., Major,J., Matthews,C.,  
 McCarthy,M., Meldrum,J., Menus,L., Mihova,T., Mlenga,V.,  
 Murphy,T., Naylor,J., Nguyen,C., Nicol,R., Norbu,C., Norman,C.H.,  
 O'Connor,T., O'Donnell,P., O'Neill,D., Oliver,J., Peterson,K.,  
 Phunhahng,P., Pierre,N., Raymond,C., Retta,R., Rise,C., Rogov,P.,  
 Roman,J., Roy,A., Schauer,S., Schupback,R., Seaman,S., Severy,P.,  
 Smith,C., Spencer,B., Stange-Thomann,N., Stojanovic,N., Talamas,J.,  
 Tesfaye,S., Theodore,J., Topham,K., Travers,M., Vassiliev,H.,  
 Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Young,G., Zainoun,J.,  
 Zembek,L., Zimmer,A. and Zody,M.  
 Direct Submission  
 Submitted (30-AUG-2002) Whitehead Institute/MIT Center for Genome  
 Research, 320 Charles Street, Cambridge, MA 02141, USA  
 All repeats were identified using RepeatMasker:  
 Smit, A.F.A. & Green, P. (1996-1997)  
 http://ftp.genome.washington.edu/RW/RepeatMasker.html

TITLE  
 JOURNAL  
 COMMENT  
 Center code: W13R  
 Center: Whitehead Institute/ MIT Center for Genome Research  
 Web site: http://www-seq.wi.mit.edu  
 Contact: sequence\_submissions@genome.wi.mit.edu  
 Project Information  
 Center project name: L28114  
 Center clone name: 452\_I\_12

NOTE: This record contains 81 individual  
 sequencing reads that have not been assembled into  
 contigs. Runs of N are used to separate the reads  
 and the order in which they appear is completely  
 arbitrary. Low-pass sequence sampling is useful for  
 identifying clones that may be gene-rich and allows  
 overlap relationships among clones to be deduced.  
 However, it should not be assumed that this clone  
 will be sequenced to completion. In the event that  
 the record is updated, the accession number will  
 be preserved.

```

1 688: contig of 688 bp in length
* 689 788: gap of 100 bp
* 789 1492: contig of 704 bp in length
* 1493 1592: gap of 100 bp
* 1593 2304: contig of 712 bp in length
* 2305 2404: gap of 100 bp
* 2405 3111: contig of 707 bp in length
* 3112 3211: gap of 100 bp
* 3212 3928: contig of 717 bp in length
* 3929 4029: gap of 100 bp
* 4029 4758: contig of 730 bp in length
* 4758 4858: gap of 100 bp

```

```
* 4859 5578: contig of 720 bp in length
* 5579 5578: gap of 100 bp
* 5679 6395: contig of 717 bp in length
* 6396 6495: gap of 100 bp
* 6496 7212: contig of 717 bp in length
* 7213 7312: gap of 100 bp
* 7313 8028: contig of 716 bp in length
* 8029 8128: gap of 100 bp
* 8129 8857: contig of 723 bp in length
* 8858 9662: contig of 705 bp in length
* 9663 9762: gap of 100 bp
* 9763 10469: contig of 707 bp in length
* 10470 10569: gap of 100 bp
* 10570 11275: contig of 706 bp in length
* 11276 11375: gap of 100 bp
* 11376 12101: contig of 726 bp in length
* 12102 12201: gap of 100 bp
* 12202 12918: contig of 717 bp in length
* 12919 13018: gap of 100 bp
* 13019 13710: contig of 692 bp in length
* 13711 13810: gap of 100 bp
* 13811 14524: contig of 714 bp in length
* 14525 15348: gap of 100 bp
* 15349 15448: gap of 100 bp
* 15449 16130: contig of 682 bp in length
* 16131 16230: gap of 100 bp
* 16231 16904: contig of 674 bp in length
* 16905 17004: gap of 100 bp
* 17005 17701: contig of 697 bp in length
* 17702 18518: contig of 717 bp in length
* 18519 18618: gap of 100 bp
* 18619 19331: contig of 713 bp in length
* 19332 19431: gap of 100 bp
* 19432 20131: contig of 700 bp in length
* 20132 20933: contig of 702 bp in length
* 20934 21033: gap of 100 bp
* 21034 21743: contig of 710 bp in length
* 21744 21843: gap of 100 bp
* 21844 22563: contig of 720 bp in length
* 22564 23381: gap of 100 bp
* 23382 23481: contig of 716 bp in length
* 23482 24183: gap of 100 bp
* 24184 24283: gap of 100 bp
* 24284 24979: contig of 696 bp in length
* 24980 25079: gap of 100 bp
* 25080 25806: contig of 727 bp in length
* 25807 25906: gap of 100 bp
* 25907 26601: contig of 695 bp in length
* 26602 27416: contig of 715 bp in length
* 27417 27516: gap of 100 bp
* 27517 28224: contig of 708 bp in length
* 28225 28324: gap of 100 bp
* 28325 29023: contig of 699 bp in length
* 29024 29123: gap of 100 bp
* 29124 29839: contig of 716 bp in length
* 29840 29939: gap of 100 bp
* 29940 30650: contig of 711 bp in length
* 30651 30750: gap of 100 bp
* 30751 31462: contig of 712 bp in length
* 31463 32274: gap of 100 bp
* 32275 32374: contig of 712 bp in length
* 32375 33093: gap of 100 bp
* 33094 33193: gap of 100 bp
* 33194 33910: contig of 717 bp in length
* 33911 34010: gap of 100 bp
* 34011 34721: contig of 711 bp in length
```

```
* 34722 34821: gap of 100 bp
* 34822 35545: contig of 724 bp in length
* 35546 35645: gap of 100 bp
* 35646 36356: contig of 711 bp in length
* 36357 36456: gap of 100 bp
* 36457 37175: contig of 719 bp in length
* 37176 37275: gap of 100 bp
* 37276 38005: contig of 730 bp in length
* 38006 38105: gap of 100 bp
* 38106 38822: contig of 717 bp in length
* 38823 39642: gap of 100 bp
* 39643 39742: contig of 720 bp in length
* 39743 40463: gap of 100 bp
* 40464 40563: gap of 100 bp
* 40564 41293: contig of 730 bp in length
* 41294 41393: gap of 100 bp
* 41394 42102: contig of 709 bp in length
* 42103 42202: gap of 100 bp
* 42203 42887: contig of 685 bp in length
* 42888 42987: gap of 100 bp
* 42989 43705: contig of 718 bp in length
* 43706 43805: gap of 100 bp
* 43806 44515: contig of 711 bp in length
* 44516 44615: gap of 100 bp
* 44617 45327: contig of 711 bp in length
* 45328 45427: gap of 100 bp
* 45428 46150: contig of 723 bp in length
* 46151 46250: gap of 100 bp
* 46251 46973: contig of 723 bp in length
* 46974 47073: gap of 100 bp
* 47074 47783: contig of 709 bp in length
* 47783 47882: gap of 100 bp
* 47883 48595: contig of 713 bp in length
* 48596 48695: gap of 100 bp
* 48696 49391: contig of 696 bp in length
* 49392 49491: gap of 100 bp
* 49492 50213: contig of 722 bp in length
* 50214 50313: gap of 100 bp
* 50314 51002: contig of 689 bp in length
* 51003 51102: gap of 100 bp
* 51103 51814: contig of 712 bp in length
* 51815 51914: gap of 100 bp
* 51915 52612: contig of 698 bp in length
* 52613 52712: gap of 100 bp
* 52713 53430: contig of 718 bp in length
* 53431 53530: gap of 100 bp
* 53531 54239: contig of 709 bp in length
* 54240 54339: gap of 100 bp
* 54340 55071: contig of 732 bp in length
* 55072 55171: gap of 100 bp
* 55172 55901: contig of 730 bp in length
* 55902 56001: gap of 100 bp
* 56002 56722: contig of 721 bp in length
* 56723 56822: gap of 100 bp
* 56823 57533: contig of 711 bp in length

Query Match      37.4%; Score 156; DB 2; Length 65608;
Beet Local Similarity 87.2%; Pred. No. 2e-20;
Matches 171; Conservative 0; Mismatches 25; Indels 0; Gaps 0;
```

12 GTGGCCGAGGACCTCCGACCTCAGGCACTGAGAGAGACATGTGGGAGATGATGA 71  
13521 GTGGCGGAGAGAGTCCGACCTCAGGCACTGAGAGAGAGAGAGTGGGAGATGGA 13462

QY 72 GAGCCAGCCGACAGAGATGCTAGCCTGATCTGCCAGCCGCTTACCCCTACTGCG 131  
DB 13461 GAGCCACTTCCACGAGATGATACATGATCTGCCCCAGCCGCTGATCATGCG 13402

QY 132 CCAGGCCAAGTGTCTGCTCCAGCTGACAGAGAGAGAGTGTGCACAGCCCGGCTCAC 191  
DB 13401 CCAGGCCAAGTGTGTATCCAACTGGGAGAGAGAGAGTGTGTGCACAGCCCGGCTCAC 13342







```

FT      /tag= b
FT      /note= "Encodes caspase recruitment domain (CARD).
FT      Specifically claimed in claim 2"
XX      US2002164703-A1.
XX      07-NOV-2002.
XX      19-DEC-2001; 2001US-0032159.
XX      21-DEC-2000; 2000US-257457P.
XX      (PAWL/) PAWLOWSKI K.
XX      (REED/) REED J C.
XX      (GODZ/) GODZIK A.
XX      PAWLOWSKI K, Reed JC, Godzik A;
XX      WPI, 2003-288137/28.
XX      P-PSDB; ABG76062.
XX      New isolated CARD-containing nucleic acids, useful for the diagnosis
XX      and treatment of disorders with aberrant expression or activity of the
XX      CARD-containing polypeptide, such as cancer, stroke, arthritis, heart
XX      failure and AIDS.
XX      Claim 1; Fig 3; 34pp; English.
XX      The invention relates to an isolated nucleic acid molecule encoding a
XX      caspase recruitment domain (CARD) containing polypeptide. CARD containing
XX      polypeptides are involved in apoptosis (as caspase activators and caspase
XX      inhibitors), cell adhesion, inflammation and cytokine receptor
XX      signaling. The methods and compositions of the present invention are
XX      useful for the diagnosis and treatment of disorders associated with the
XX      aberrant expression or activity of the CARD containing polypeptide such
XX      as cancer, glioma, lymphoma, adenocarcinoma, sarcoma, melanoma, keloid,
XX      hamartoma, leukoemia, lymphoma, keratinocyte hyperplasia, neoplasia,
XX      benign prostatic hypertrophy, inflammatory diseases such as arthritis, lupus,
XX      Sjogren's syndrome, Crohn's disease, ulcerative colitis, graft versus
XX      host disease, sepsis, abnormal cell death diseases such as stroke,
XX      myocardial infarction, heart failure, neurodegenerative diseases like
XX      Parkinson's disease and Alzheimer's disease, and HIV infection. The
XX      present sequence represents cDNA encoding the human caspase recruitment
XX      domain containing protein, CARD-12X.
XX      Sequence 417 BP; 84 A; 131 C; 128 G; 72 T; 2 other;
SQ
Query Match          99.5%; Score 415; DB 25; Length 417;
Best Local Similarity 100.0%; Pred. No. 5.4e-83;
Matches 415; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Db      301 ACCCTGATACCCGGGCTGACGCTGATGTTGACTTACGTAACCTTACCGGTAGAGCTCC 360
QY      361 GACTTGCACGGTTTGGAGGCACTTCTAGAACCTCAGGCTCTGTGTAACCCAG 415
Db      361 GACTTGCACGGTTTGGAGGCACTTCTAGAACCTCAGGCTCTGTGTAACCCAG 415

RESULT 2
AAD13448
ID      AAD13448 standard; DNA; 3417 BP.
XX
XX      AAD13448;
XX
XX      06-NOV-2001 (first entry)
XX
XX      Human caspase recruitment domain-14 (CARD-14) genomic DNA.
XX
XX      Human; caspase recruitment domain-14; CARD-14; chromosome 17;
XX      nuclear factor-kappa B; NF-kB; cell growth; cell death; cancer; therapy;
XX      autoimmune disorder; systemic lupus erythematosus; neurological disorder;
XX      Alzheimer's disease; Parkinson's disease; inflammatory disorder; anaemia;
XX      haematological disorder; myelodysplastic syndrome; myocardial infarction;
XX      stroke; immune disorder; Crohn's disease; allergic rhinitis; infection;
XX      cell signalling disorder; cytostatic; immunosuppressive; nootropic;
XX      neuroprotective; antiviral; antibacterial; ds.
XX
XX      Homo sapiens.
XX
XX      Location/Qualifiers
XX      FH      Key
XX      FT      CDS
XX      FT      1..3417
XX      FT      /tag= a
XX      FT      /product= "Human caspase recruitment domain-14"
XX      FT      /transl_except= (pos:2098..2100, aa:Cys)
XX      FT      /note= "CDS does not include stop codon"
XX      FT      /partial
XX      FT      1..211
XX      FT      /tag= b
XX      FT      /number= "1"
XX      FT      212..415
XX      FT      /tag= c
XX      FT      /number= "2"
XX      FT      416..741
XX      FT      /tag= d
XX      FT      /number= "3"
XX      FT      742..909
XX      FT      /tag= e
XX      FT      /number= "4"
XX      FT      910..1029
XX      FT      /tag= f
XX      FT      /number= "5"
XX      FT      1030..1209
XX      FT      /tag= g
XX      FT      /number= "6"
XX      FT      1210..1359
XX      FT      /tag= h
XX      FT      /number= "7"
XX      FT      1360..1476
XX      FT      /tag= i
XX      FT      /number= "8"
XX      FT      1477..1619
XX      FT      /tag= j
XX      FT      /number= "9"
XX      FT      1620..1714
XX      FT      /tag= k
XX      FT      /number= "10"
XX      FT      1715..1778
XX      FT      /tag= l
XX      FT      /number= "11"
XX      FT      1779..2001
XX      FT      /tag= m
XX      FT      /number= "12"
XX      FT      2002..2128
XX      FT      exon

```

FT FT /tag= n  
FT exon /number= "13"  
FT 2129..2369  
XX /tag= o  
FT exon /number= "14"  
FT 2370..2433  
FT /tag= p  
FT exon /number= "15"  
FT 2434..2548  
FT /tag= q  
FT exon /number= "16"  
FT 2549..2719  
FT /tag= r  
FT exon /number= "17"  
FT 2720..2841  
FT /tag= s  
FT exon /number= "18"  
FT 2842..2957  
FT /tag= t  
FT exon /number= "19"  
FT 2958..3138  
FT /tag= u  
FT exon /number= "20"  
FT 3139..33417  
FT /tag= v  
PN /number= "21"  
PN WO200159065-A2.  
PD 16-AUG-2001.  
XX 22-JAN-2001; 2001WO-USO2087.  
XX P-PSDB; AAB07165.  
XX 09-FEB-2000; 2000US-0181159.  
XX (MILL-) MILLENNIUM PHARM INC.  
PA Bertin J;  
PI DR WPI; 2001-497073/54.  
XX P-PSDB; AAB07165.  
XX An isolated caspase recruitment domain polypeptide useful for  
PT regulating growth and cell death and useful for the treatment of cancer  
PT -  
XX Disclosure; Fig 2A-2C; 109pp; English.

The present sequence is human caspase recruitment domain-14 (CARD-14) DNA. The CARD-14 gene is located on chromosome 17. The CARD-14 is used for the detection of modulators that modulates the ability of CARD-14 to bind to Bcl-10 and stimulates phosphorylation of Bcl-10 or activation of nuclear factor-kappa B (NF-KB). The CARD-14 is useful for regulating growth and cell death and useful for the treatment of cancer. It is also useful for the treatment of autoimmune disorders (e.g., systemic lupus erythematosus), neurological disorders e.g., Alzheimer's and Parkinson's disease, inflammatory disorders, haematological disorders (e.g., anaemia, myelodysplastic syndromes), myocardial infarctions, strokes, immune disorders (e.g., Crohn's disease, allergic rhinitis), cell signalling disorders and certain viral and bacterial infections.

Sequence 3417 BP; 739 A; 1039 C; 1102 G; 537 T; 0 other;

```

Query Match Similarity 99.5%; Score 415; DB 22; Length 3417;
Best Local Similarity 100.0%; Pred. No. 7.2e-83;
Matches 415; Conservative 0; Mismatches 0; Indels 0; Gaps 0

QY 1 ATGGGGGAACCTGTGCCAGAGGACTCCGACCTCACGGCACTGGACGAGAGGACACTGTGG 60
DB 1 ATGGGGGAACCTGTGCCAGAGGACTCCGACCTCACGGCACTGGACGAGAGGACACTGTGG 60
QY 61 GAGATGATGAGAGGCAACCGCCACACAGATCTTACGTGATTTGCCCAAGCCGCTCACC 120

```

Db	61	GAGATATGGAAGGCAACGGCAACAGATGCTAGCTGGCACTTGGCCCAAGCCGCTCAC	120
Qy	121	CCCTTACTCCGCCAAGCCCAAGTGTGTGTCACGCTGGAACGAGAGAGTGTCTGCACAG	180
Db	121	CCCTTACTCCGCCAAGCCCAAGTGTGTGTCACGCTGGAACGAGAGAGAGTGTCTGCACAG	180
Qy	181	CCCCGCTCACCACAGAGCGCATGCGGGCCGGGCACTTGTGATTTGTGAAAGACTCGA	240
Db	181	CCCCGCTCACCACAGAGCGCATGCGGGCCGGGCACTTGTGATTTGTGAAAGACTCGA	240
Qy	241	GGGAAAGAGGGGCCATGCTTCTCTGGAAGCCTGAAGTTCCACAACCTGAAGCTTAC	300
Db	241	GGGAAAGAGGGGCCATGCTTCTCTGGAAGCCTGAAGTTCCACAACCTGAAGCTTAC	300
Qy	301	ACCCTGTCAACGGGCTGACAGCTGATTTTACTTACCTGAATACCTTTAGCGGTGAAGCTCC	360
Db	301	ACCCTGTCAACGGGCTGACAGCTGATTTTACTTACCTGAATACCTTTAGCGGTGAAGCTCC	360
Qy	361	GACTTGAAGGTTTGGCAGCAGCACTTCTTGAAGAACTCAGGCTCCCGTGAACCCGAG	415
Db	361	GACTTGAAGGTTTGGCAGCAGCACTTCTTGAAGAACTCAGGCTCCCGTGAACCCGAG	415

```
RESULT 3
AAD13447
ID AAD13447 standard; cDNA; 3931 BP.
```

AC AAD13447

DT 06-NOV-2001 (first entry)

DE Human caspase recruitment domain-14 (CARD-14) cDNA

KM Human caspase recruitment domain-14; CARD-14; chromosome 17;  
KM nuclear factor-kappa B; NF- $\kappa$ B; cell growth; cell death; cancer; therapy;  
KM autoimmune disorder; systemic lupus erythematosus; neurological disorder;  
KM Alzheimer's disease; Parkinson's disease; inflammatory disorder; anemia;  
KM haematological disorder; myelodysplastic syndrome; myocardial infarction;  
KM stroke; immune disorder; Crohn's disease; allergic rhinitis; infection;  
KM cell signalling disorder; cytostatic; immunosuppressive; nocrotropic;  
KM neuroprotective; antiviral; antibacterial; ss.

OS Homo sapiens.

	Key	Location/Qualifiers
FH	CDS	207..3221
FT		

FT /product= "Human caspase recruitment domain-14"  
FT /note= "Residues 207-3218 is specifically claimed  
FT as SEQ ID NO 3 in claim 9 of the specification"

PN W0200159065-A2.

PD 16-AUG-2001

PF 22-JAN-2001; 2001WO-US02087.

PR 09-FEB-2000; 2000US-0181159.

PA (MILL-) MILLENNIUM PHARM INC.

PI Bertin J;

DR WPI; 2001-497073/54.


 文部科学省

PT regulating growth and cell death and useful for the treatment of cancer

•

100

CC The present sequence is human caspase recruitment domain-14 (CARD-14)

CC CDNA. The CARD-14 gene is located on chromosome 17. The CARD-14 is used  
 CC for the detection of modulators that modulates the ability of CARD-14 to  
 CC bind to Bcl-10 and stimulates phosphorylation of Bcl-10 or activation  
 CC of nuclear factor-kappa B (NF-kB). The CARD-14 is useful for regulating  
 CC growth and cell death and useful for the treatment of cancer. It is  
 CC also useful for the treatment of autoimmune disorders (e.g., systemic  
 CC lupus erythematosus), neurological disorders (e.g., Alzheimer's and  
 CC Parkinson's disease), inflammatory disorders, haematological disorders  
 CC (e.g., anaemia, myelodysplastic syndromes), myocardial infarctions,  
 CC strokes, immune disorders (e.g., Crohn's disease, allergic rhinitis),  
 CC cell signalling disorders and certain viral and bacterial infections.

XX Sequence 3931 BP, 799 A, 1187 C, 1263 G, 662 T, 0 other;

Query Match 84.2%; Score 351; DB 22; Length 3931;  
 Best Local Similarity 100.0%; Pred. No. 1.2e-68;  
 Matches 351; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGGGGAACTGTGCGGAGGAGTCCGCACTCAACGCGACATGACGAGAGACACTGTGG 60  
 DB 207 ATGGGGAACTGTGCGGAGGAGTCCGCACTCAACGCGACATGACGAGAGACACTGTGG 266  
 QY 61 GAGATGATGAGAGGACCGCCACAGATGCTAGCGCTGATCTGCCAGCCGCTCAAC 120  
 DB 267 GAGATGATGAGAGGACCGCCACAGATGCTAGCGCTGATCTGCCAGCCGCTCAAC 326  
 QY 121 CCTTACCTGCGGCGAGGCGCAAGTGTCTGCGACGCTGACGAGAGAGAGTGTCTCAACG 180  
 DB 327 CCTTACCTGCGGCGAGGCGCAAGTGTCTGCGACGCTGACGAGAGAGAGTGTCTCAACG 386  
 QY 181 CCCCGGCTCACCAACAGCGCCATGCGGAGCGGCGACCTTGCTGATTTGCTGAAGACTGA 240  
 DB 387 CCCCGGCTCACCAACAGCGCCATGCGGAGCGGCGACCTTGCTGATTTGCTGAAGACTGA 446  
 QY 241 GGGAAAGACGGGGCCATCGCTTCTGTGAGAGCCTGAAGTTCCAAACCTGACGTCTAC 300  
 DB 447 GGGAAAGACGGGGCCATCGCTTCTGTGAGAGCCTGAAGTTCCAAACCTGACGTCTAC 506  
 QY 301 ACCCTGCTCACCGGCGCTGACGCTGATGCTGATCTTCACTTAAGCGGT 351  
 DB 507 ACCCTGCTCACCGGCGCTGACGCTGATGCTGATCTTCACTTAAGCGGT 557

## RESULT 4

AAS98203 standard; cDNA, 3948 BP.

AC AAS98203;

DT 12-MAR-2002 (first entry)

DE DNA encoding plakoglobin interacting protein #3.

XX Human; plakoglobin; cytosolic; osteopathic; dermatological; cardiac;  
 KM plakoglobin related disease; skin carcinoma; acantholytic disease;  
 KM basal cell carcinoma; squamous cell carcinoma; Naxos disease; PCR primer;  
 KM extramammary Paget's disease; heart disease; skin blistering;  
 KM subcorneal acantholysis; Grover's disease; Hailey-Hailey's disease;  
 KM Darier's disease; ectodermal dysplasia; skin fragility syndrome; ss.

OS Homo sapiens.

XX WO200185933-A2.

XX 15-NOV-2001.

XX 02-MAY-2001; 2001WO-EP04872.

XX 09-MAY-2000; 2000EP-0201668.

XX (VLA-) VLAAMS INTERUNIVERSITAIR INST BIOTECHNOG.

XX Van Roy F, Bonne S, Vanlandeschoot A;  
 XX PI

XX WPI, 2002-062246/08.  
 DR P-PSDB; AAU73247.

XX New polypeptide, useful for treating skin carcinoma or acantholytic  
 PT disease such as Grover's and Darier's disease, comprises a protein  
 PT interacting with human plakoglobin and involved in transduction of  
 PT plakoglobin related signal to nucleus -

XX Claim 7; Figure 3; 98pp; English.

CC The invention relates to an isolated plakoglobin interacting polypeptide  
 CC (1). (1) is useful as a medicament and in the manufacture of a  
 CC medicament for treating plakoglobin related diseases, such as skin  
 CC carcinoma or an acantholytic disease, and to screen compounds that  
 CC interfere with the interaction of the polypeptide with plakoglobin  
 CC The plakoglobin related diseases include basal cell carcinoma, squamous  
 CC cell carcinoma, extramammary Paget's disease, Naxos disease, heart  
 CC diseases, skin blistering and acantholytic diseases such as subcorneal  
 CC acantholysis, Grover's disease, Hailey-Hailey's disease or Darier's  
 CC disease, and ectodermal dysplasia/skin fragility syndrome. AAS98201-  
 CC AAS98286 represent novel human plakoglobin interacting protein  
 CC coding sequences and PCR primers of the invention.

XX Sequence 3948 BP, 717 A, 1218 C, 1327 G, 686 T, 0 other;

Query Match 28.2%; Score 117.8; DB 24; Length 3948;  
 Best Local Similarity 59.3%; Pred. No. 6.2e-17;  
 Matches 200; Conservative 0; Mismatches 137; Indels 0; Gaps 0;

QY 18 CAGGAGCTCCGCACTCAACGCGCACTGAGAGAGACACTGTGGAGATGAGAGAGCA 77  
 DB 86 CAGGAGCTCCGCACTCAACGCGCACTGAGAGAGACACTGTGGAGATGAGAGAGCA 145  
 QY 78 CCGCAACAGATGATCGCTGATCTGCCAGCCGCTCAACCCCTTACCTGCGCAGGC 137  
 DB 146 CCGCAACAGATGATCGCTGATCTGCCAGCCGCTCAACCCCTTACCTGCGCAGGC 205  
 QY 138 CAAAGTGTCTGTGCGACGCTGAGAGAGAGAGTGTCTGACAGCCCGGCTCAACAG 197  
 DB 206 CCGAGTGTCTGTGCGACGCTGAGAGAGAGAGTGTCTGACAGCCCGGCTCAACAG 265  
 QY 198 CGCGATGCGGGCGGCGCACTTGTGATTTGCTGAAGATCGAGGGAGAGCGGGCCAT 257  
 DB 266 CGTCAACCGACCGGGCGGCTGATGACATCTTGCGCTGCCGTGGAGAGGGGCTATGA 325  
 QY 258 CGCCTTCTGAGAGGCTGAGAGTTCACCAACCTTGACCTTACACCTTGATCAACCGGCT 317  
 DB 326 GGCCTTCTGAGAGGCTGAGAGTTCACCAACCTTGACCTTACACCTTGATCAACCGGCT 385  
 QY 318 GCAGCTGATGTTGACTTCACTTAAGCGGTGAG 354  
 DB 386 GCAACCGCGCGAGCGCTGCTCATGATCTTCATGAG 422

## RESULT 5

AAS05388 standard; cDNA, 3949 BP.

AC AAS05388;

DT 12-SEP-2001 (first entry)

DE Human caspase recruitment domain, CARD-10 cDNA sequence.

XX Human; caspase recruitment domain; CARD-10; Bcl-10; NF-kappaB;  
 KM apoptosis; hyperproliferative disorder; autoimmune; neurological;  
 KM inflammatory disorder; viral infection; stress-related response; ss.

OS Homo sapiens.

XX Key Location/Qualifiers  
 XX CDS 41..3139

```

FT      /*tag= a
FT      /product= "CARD-10"
FT      /note= "ORF is specifically claimed"
XX
XX      MO200140468-A2.
XX
XX      07-JUN-2001.
XX
XX      01-DEC-2000; 2000WO-US32716.
XX
XX      03-DEC-1999; 99US-0168780.
XX      18-FEB-2000; 2000US-0507533.
XX      25-FEB-2000; 2000US-0513904.
XX      10-OCT-2000; 2000US-0685791.
XX
XX      (MILL-) MILLENNIUM PHARM INC.
XX
XX      Bertin J;
XX
XX      WPI; 2001-367809/38.
XX      P-PSDB; AAU01206.
XX
XX      Novel caspase recruitment domain (CARD) proteins, CARD-9, CARD-10,
XX      CARD-11, useful as targets for therapy, as immunogens, and in screening
XX      and detection assays.
XX
XX      Claim 2; Fig 10A-10C; 145bp; English.
XX
XX      The present sequence encoding for novel human caspase recruitment
XX      domain, CARD-10 is isolated from a human skin cDNA library.
XX      Also described are novel human sequences for CARD-9 and CARD-11
XX      (AAU01205, AAU01207) and rat CARD-9 (AAU01204). CARD-9, CARD-10 and
XX      CARD-11 interact with Bcl-10 which is thought to activate nuclear factor
XX      (NF)-kappaB and apoptosis. The sequences of the invention can be used for
XX      treating a disorder associated with abnormal levels of apoptosis by
XX      modulating the expression or activity of CARD-9, CARD-10, or CARD-11.
XX      They can be used for the treatment of hyperproliferative disorders
XX      (e.g. cancer), autoimmune disorders (e.g. systemic lupus erythematosus),
XX      neurological disorders (e.g. Alzheimer's disease), inflammatory disorders
XX      (e.g. Crohn's disease), and viral infection (e.g. HIV). The CARD
XX      polypeptide, polynucleotide and an antibody which selectively binds to
XX      CARD can be used in screening and detection assays (e.g. chromosomal
XX      mapping, tissue typing), predictive medicine (prognostic assays,
XX      monitoring clinical trials, and therapy (treatment and prophylaxis). The
XX      CARD polypeptide may be used to screen for drugs that bind to and/or
XX      modulate it. CARD sequences are potential targets for regulating
XX      inflammation, cancer, NF-kappaB signaling, stress-related response and
XX      apoptosis in human disease. A host cell containing a polynucleotide
XX      encoding CARD can be used to create transgenic animals.
XX
XX      Sequence 3949 BP; 724 A; 1222 C; 1319 G; 684 T; 0 other;
XX
XX      Query Match      28.2%; Score 117.8; DB 22; Length 3949;
XX      Best Local Similarity 59.3%; Pred. No. 6.2e-17;
XX      Matches 200; Conservative 0; Mismatches 137; Indels 0; Gaps 0;

```

```

DB      322 GGGCTTCCTGAGACCCCTGGAGTCTCACTACCCCGAACACTTCAGCTGCTACAGGCCA 381
OY      318 GCAGCTGATGTGACTTCAGTAACCTTAGCGGTGAG 354
DB      382 GGAACCCGCGCCGCTGCTCAGATGATCTCGATGAG 418

RESULT 6
ABA00333
ID      ABA00333 standard; cDNA; 3949 BP.
XX
XX      ABA00333;
XX
XX      09-DEC-2002 (first entry)
XX
XX      Human CARD-10 cDNA.
XX
XX      Gene; rat; human; caspase recruitment domain; CARD-9; CARD-10;
XX      CARD-11; apoptosis; inflammation; cell growth; cell death;
XX      lymphocyte activation; cancer; melanoma; autoimmune disease;
XX      arthritis; neurological disorder; Alzheimer's disease; ss.
XX
XX      Homo sapiens.
XX
XX      Key      Location/Qualifiers
XX      CDS      41..3139
XX      FT      /*tag= a
XX      FT      /product= "CARD-10"
XX
XX      MO200270652-A2.
XX
XX      12-SEP-2002.
XX
XX      28-FEB-2002; 2002WO-US06147.
XX
XX      02-WAR-2001; 2001US-0798412.
XX
XX      (MILL-) MILLENNIUM PHARM INC.
XX
XX      Bertin J;
XX
XX      WPI; 2002-698749/75.
XX      P-PSDB; AAG79554.
XX
XX      CARD-9, CARD-10 or CARD-11 polypeptides and polynucleotides, useful for
XX      treating disorders associated with inappropriate apoptosis or
XX      lymphocyte activation, e.g. cancer.
XX
XX      Claim 5; Fig 10; 151bp; English.
XX
XX      This sequence encodes human caspase recruitment domain (CARD)-10.
XX      CARD proteins play roles in apoptotic and inflammatory signaling
XX      pathways. CARD-9, -10 and -11 participate in the network of
XX      interactions that modulate caspase activity. They are thought to be
XX      useful as modulating agents for regulating a variety of cellular
XX      processes including cell growth and cell death. CARD proteins and
XX      nucleic acids are useful for treating a disorder associated with
XX      inappropriate apoptosis or lymphocyte activation or for diagnosing
XX      subjects having or that are at risk of developing a disorder associated
XX      with aberrant CARD-9, CARD-10 or CARD-11 expression or activity, such
XX      as cancer e.g. melanoma, autoimmune disorders e.g. arthritis, or
XX      neurological disorders e.g. Alzheimer's disease.
XX
XX      Sequence 3949 BP; 724 A; 1222 C; 1319 G; 684 T; 0 other;
XX
XX      Query Match      28.2%; Score 117.8; DB 24; Length 3949;
XX      Best Local Similarity 59.3%; Pred. No. 6.2e-17;
XX      Matches 200; Conservative 0; Mismatches 137; Indels 0; Gaps 0;

```

QY 78 CCGGCAAGATGATGTAAGCTGATCTGCGCCAGCGGCTCAACCCCTACCTGGGCGAGGC 137  
 DB 142 CCGGATATGGCTGCTGCTCGCCCTGAACCCGCGCAAGCTCAAGCGCTGATCGCCAGTG 201  
 QY 138 CAAAGTGTCTGTCCAGCTGACGACGAGAGAGTGTCTGACAGCCCGGCTCAACCAAG 197  
 DB 202 CCGGCTATCAACGACGACGACGAGAGAGTGTCTGACAGCCCTTCCTCCGTCGG 261  
 QY 198 CGCCATGCGGGCGCGGACCTTGTCTGATTTGCTGAAGCTTGAGAGGAAAGCGGCGCAT 257  
 DB 262 CGTCAACCGCAACCGGCGCGCTGATGACATCTTGCGCTGCGGTGCAAGAGGCGCTATGA 321  
 QY 258 CGCCTTCTGTGAGAGAGCTGAAGTTTCAACACCTGACCTTACACCTTGGTCAACGGGCT 317  
 DB 322 GGCCTTCTGTGAGAGAGCTGAGATTCTACTACCCGACACCTCAACGCTGCTCAAGGCGCA 381  
 QY 318 GCAGCTGATGTTGACTTCACTTCACTTAAAGGCGTGAAG 354  
 DB 382 GGAACCCGCGCAGCGCTGCTCCATGATCTCGATGAG 418

RESULT 7  
 ABX11449  
 ID ABX11449 standard; cDNA, 1141 BP.  
 XX  
 AC ABX11449;  
 XX  
 DT 09-MAY-2003 (first entry)  
 XX  
 DE Human CARD-12X expressed sequence tag, GI10316320.  
 XX  
 DE Human; ss; CARD; EST; caspase recruitment domain; apoptosis; lupus;  
 KW cell adhesion; inflammation; cytokine receptor signaling; glioma;  
 KW carcinoma; adenocarcinoma; hamartoma; leukemia; lymphoma; melanoma;  
 KW CARD-containing polypeptide associated disorder; sarcoma; neoplasia;  
 KW keratinocyte hyperplasia; keloid; benign prostatic hyperplasia; sepsis;  
 KW inflammatory hyperplasia; fibrosis; restenosis; allergy; arthritis;  
 KW Sjogren's syndrome; Crohn's disease; ulcerative colitis; stroke; cancer;  
 KW graft versus host disease; abnormal cell death disease; heart failure;  
 KW myocardial infarction; neurodegenerative disease; Parkinson's disease;  
 KW Alzheimer's disease; HIV; CARD-12X; caspase activator; caspase inhibitor;  
 KW expressed sequence tag.  
 XX  
 OS Homo sapiens.  
 XX  
 PN US2002164703-A1.  
 XX  
 PD 07-NOV-2002.  
 XX  
 PF 19-DEC-2001; 2001US-0032159.  
 XX  
 PR 21-DEC-2000; 2000US-257457P.  
 XX  
 PA (PAWL/) PAWLOWSKI K.  
 PA (REED/) REED J C.  
 PA (GODZ/) GODZIK A.  
 XX  
 PI Pawlowski K, Reed JC, Godzik A;  
 XX  
 DR WPI; 2003-288137/28.  
 XX  
 PT New isolated CARD-containing nucleic acids, useful for the diagnosis  
 PT and treatment of disorders with aberrant expression or activity of the  
 PT CARD-containing polypeptide, such as cancer, stroke, arthritis, heart  
 PT failure and AIDS -  
 XX  
 PS Disclosure; Fig 7; 34pp; English.  
 CC The invention relates to an isolated nucleic acid molecule encoding a  
 CC caspase recruitment domain (CARD) containing polypeptide. CARD containing  
 CC polypeptides are involved in apoptosis (as caspase activators and caspase  
 CC inhibitors), cell adhesion, inflammation and cytokine receptor

CC signalling. The methods and compositions of the present invention are  
 CC useful for the diagnosis and treatment of disorders associated with the  
 CC aberrant expression or activity of the CARD containing polypeptide such  
 CC as cancer, glioma, carcinoma, adenocarcinoma, sarcoma, melanoma, keloid,  
 CC hamartoma, leukemia, lymphoma, keratinocyte hyperplasia, neoplasia,  
 CC benign prostatic hyperplasia, inflammatory hyperplasia, fibrosis,  
 CC restenosis, allergies, inflammatory diseases such as arthritis, lupus,  
 CC Sjogren's syndrome, Crohn's disease, ulcerative colitis, graft versus  
 CC host disease, sepsis, abnormal cell death diseases such as stroke,  
 CC myocardial infarction, heart failure, neurodegenerative diseases like  
 CC Parkinson's disease and Alzheimer's disease, and HIV infection. The  
 CC present sequence represents the human caspase recruitment domain  
 CC containing protein, CARD-12X expressed sequence tag, GI10316320.  
 XX  
 SQ Sequence 1141 BP; 318 A; 351 C; 320 G; 152 T; 0 other;  
 Query Match 26.8%; Score 111.8; DB 25; Length 1141;  
 Best Local Similarity 94.3%; Pred. No. 1.1e-15;  
 Matches 116; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 1 ATGCGGGAAGTGTGCGGACGAGGACTCCGACTCAGCGGCACTGGACGAGAGACACTGTGG 60  
 DB 161 ATGCGGGAAGTGTGCGGACGAGGACTCCGACTCAGCGGCACTGGACGAGAGACACTGTGG 220  
 QY 61 GAGATGATGAGAGAGCCACCGCCACAGGATCGTACGCTGATGCCCCCGCGCTCACCC 120  
 DB 221 GAGATGATGAGAGAGCCACCGCCACAGGATCGTACGCTGATGCCCCCGAGACACTTAA 280  
 QY 121 CCC 123  
 DB 281 CCC 283

RESULT 8  
 ABX11430  
 ID ABX11430 standard; cDNA, 3744 BP.  
 XX  
 AC ABX11430;  
 XX  
 DT 09-MAY-2003 (first entry)  
 XX  
 DE Human caspase recruitment domain containing protein, CARD-11X, cDNA.  
 XX  
 DE Human; ss; CARD; EST; caspase recruitment domain; apoptosis; cell adhesion; inflammation;  
 KW cytokine receptor signaling; cancer; glioma; carcinoma; adenocarcinoma;  
 KW CARD-containing polypeptide associated disorder; sarcoma; melanoma; ss;  
 KW hamartoma; leukemia; lymphoma; keratinocyte hyperplasia; neoplasia;  
 KW keloid; benign prostatic hyperplasia; inflammatory hyperplasia; fibrosis;  
 KW restenosis; allergy; arthritis; lupus; Sjogren's syndrome; sepsis; gene;  
 KW Crohn's disease; ulcerative colitis; graft versus host disease; stroke;  
 KW abnormal cell death disease; myocardial infarction; heart failure; human;  
 KW neurodegenerative disease; Parkinson's disease; Alzheimer's disease; HIV;  
 KW CARD-11X; caspase activator; caspase inhibitor.  
 XX  
 OS Homo sapiens.  
 XX  
 PN US2002164703-A1.  
 XX  
 PD 07-NOV-2002.  
 XX  
 PF 19-DEC-2001; 2001US-0032159.  
 XX  
 PR 21-DEC-2000; 2000US-257457P.  
 XX  
 PA (PAWL/) PAWLOWSKI K.  
 PA (REED/) REED J C.  
 PA (GODZ/) GODZIK A.  
 XX  
 PI Pawlowski K, Reed JC, Godzik A;  
 XX  
 DR WPI; 2003-288137/28.  
 XX  
 PT New isolated CARD-containing nucleic acids, useful for the diagnosis  
 PT and treatment of disorders with aberrant expression or activity of the  
 PT CARD-containing polypeptide, such as cancer, stroke, arthritis, heart  
 PT failure and AIDS -  
 XX  
 PS Disclosure; Fig 7; 34pp; English.  
 CC The invention relates to an isolated nucleic acid molecule encoding a  
 CC caspase recruitment domain (CARD) containing polypeptide. CARD containing  
 CC polypeptides are involved in apoptosis (as caspase activators and caspase  
 CC inhibitors), cell adhesion, inflammation and cytokine receptor





```

Db      357 GAAAGAGCGCTTGTGGAGAAATGTGAGTAAACCGGCAACAGCTCAGCGCTATATC 416
Qy      103 TGGCCAGCCGCTTACACCCCTACCTGGCCAGGCAAGGTGTGTCCAGCTGACGAG 162
Db      417 AACCTGCAAGCTCAACCCCTACTGCTGTGTAAGTGAAGTCAATTGATGAGCAGATGAA 476
Qy      163 GAGAGGTGTGTGCAACGCCCCGCTCAACAAAGCCGATGCGGGCCGCACTTGCTG 222
Db      477 GATGAAGTGTAAATGCCCCCTATGTGCTGCATCAAGATCAACGAGAGGCGGCTGTG 536
Qy      223 GATTGCTGAAGACTCGAGGGAAGAACGGGCAATGCGCTTCTGTGAGAGCTTGAAGTTC 282
Db      537 GACATTTCACTACATCCAGAGGGCAAGGGGCTATGTGTCTTCTTGAGAGGCTTGAATTT 596
Qy      283 CACAACCTGACGTGTACACCTGTGTCACGGGCGTCAAGCTGATGTTGACTTCAAGTAC 342
Db      597 TATTACCAAGAACTGTACAACTGGTGACTGGGAAAGAGCCCACTCGGAGATTCTCAC 656
Qy      343 TTTAGCGGTGAG 354
Db      657 ATTGTGTGAG 668

```

## RESULT 10

ABA00334  
ID ABA00334 standard; cDNA; 4276 BP.

XX ABA00334;

DT 09-DEC-2002 (first entry)

DE Human CARD-11 cDNA.

XX Gene; rat; human; caspase recruitment domain; CARD-9; CARD-10;  
KW CARD-11; apoptosis; inflammation; cell growth; cell death;  
KW lymphocyte activation; cancer; melanoma; autoimmune disease;  
KW arthritis; neurological disorder; Alzheimer's disease; ss.

OS Homo sapiens.

XX Key Location/Qualifiers  
FH 328..3771  
FT /\*tag=a  
FT /product="CARD-11"

PN WO200270652-A2.

PD 12-SEP-2002.

PF 28-FEB-2002; 2002WO-US06147.

PR 02-MAR-2001; 2001US-0798412.

PA (MILL-) MILLENNIUM PHARM INC.

PI Bertin J;

XX WPI; 2002-698749/75.

DR P-P8DB; AAG79555.

PT CARD-9, CARD-10 or CARD-11 polypeptides and polynucleotides, useful for  
PT treating disorders associated with inappropriate apoptosis or  
PT lymphocyte activation, e.g. cancer

PS Disclosure; Fig 14; 151pp; English.

XX This sequence encodes human caspase recruitment domain (CARD)-11.  
CC CARD proteins play roles in apoptotic and inflammatory signalling  
CC pathways. CARD-9, -10 and -11 participate in the network of  
CC interactions that modulate caspase activity. They are thought to be  
CC useful as modulating agents for regulating a variety of cellular  
CC processes including cell growth and cell death. CARD proteins and

CC nucleic acids are useful for treating a disorder associated with  
CC inappropriate apoptosis or lymphocyte activation or for diagnosing  
CC subjects having or that are at risk of developing a disorder associated  
CC with aberrant CARD-9, CARD-10 or CARD-11 expression or activity, such  
CC as cancer e.g. melanoma, autoimmune disorders e.g. arthritis, or  
CC neurological disorders e.g. Alzheimer's disease.

XX Sequence 4276 BP; 1045 A; 1256 C; 1320 G; 655 T; 0 other;

Query Match 24.9%; Score 104; DB 24; Length 4276;

Best Local Similarity 58.3%; Pred. No. 7.2e-14;

Matches 182; Conservative 0; Mismatches 130; Indels 0; Gaps 0;

Qy 43 GAGAGAGACACTGTGGAGATGATGAGAGCCACCGGACAGATCTGACGCTGCATC 102

Db 358 GAAAGAGAGCCCTTGTGGAGAGATGTGAGTGTAAACCGGCAATGTCACCCGCTATATC 417

Qy 103 TGGCCAGCCGCTTACACCCCTACCTGGCCAGGCAAGGTGTGTCCAGCTGACGAG 162

Db 418 AACCTGCAAGCTCAACCCCTACTGCTGTGTAAGTGAAGTCAATTGATGAGCAGATGAA 477

Qy 163 GAGAGGTGTGTGCAACGCCCCGCTCAACAAAGCCGATGCGGGCCGCACTTGCTG 222

Db 478 GATGAAGTGTAAATGCCCCCTATGTGCTGCATCAAGATCAACGAGAGGCGGCTGTG 537

Qy 223 GATTGCTGAAGACTCGAGGGAAGAACGGGCAATGCGCTTCTGTGAGAGCTTGAAGTTC 282

Db 538 GACATTTCACTACATCCAGAGGGCAAGGGGCTATGTGTCTTCTTGAGAGGCTTGAATTT 597

Qy 283 CACAACCTGACGTGTACACCTGTGTCACGGGCGTCAAGCTGATGTTGACTTCAAGTAC 342

Db 598 TATTACCAAGAACTGTACAACTGGTGACTGGGAAAGAGCCCACTCGGAGATTCTCAC 657

Qy 343 TTTAGCGGTGAG 354

Db 658 ATTGTGTGAG 668

## RESULT 11

AAH08620  
ID AAH08620 standard; cDNA; 765 BP.

XX AAH08620;

AC 26-JUN-2001 (first entry)

DE Human cDNA clone (5'-primer) SEQ ID NO:5455.

XX Human; primer; detection; diagnosis; antisense therapy; gene therapy; ss.

OS Homo sapiens.

PN EP1074617-A2.

PD 07-FEB-2001.

PF 28-JUL-2000; 2000EP-0116126.

PR 29-JUL-1999; 99JP-0248036.

PR 27-AUG-1999; 99JP-0300253.

PR 11-JAN-2000; 2000JP-0118776.

PR 02-MAY-2000; 2000JP-0183767.

PR 09-JUN-2000; 2000JP-0241899.

PA (HELT-) HELIX RES INST.

XX Oca T, Isegai T, Mshikawa T, Hayaeshi K, Saito K, Yamamoto J;

PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;

DR WPI; 2001-318749/34.

PT Primer sets for synthesizing polynucleotides, particularly the 5602  
PT full-length cDNAs defined in the specification, and for the detection

PT and/or diagnosis of the abnormality of the proteins encoded by the  
 PT full-length cDNAs -

PS Claim 1; SEQ ID 5455; 2537pp + CD ROM; English.

XX The present invention describes primer sets for synthesizing 5602  
 CC full-length cDNAs defined in the specification. Where a primer set  
 CC comprises: (a) an oligo-dT primer and an oligonucleotide complementary  
 CC to the complementary strand of a polynucleotide which comprises one of  
 CC the 5602 nucleotide sequences defined in the specification, where the  
 CC of an oligonucleotide comprising a sequence complementary to the  
 CC complementary strand of a polynucleotide which comprises a 5'-end  
 CC sequence and an oligonucleotide comprising a sequence complementary to a  
 CC polynucleotide which comprises a 3'-end sequence, where the  
 CC oligonucleotide comprises at least 15 nucleotides and the combination of  
 CC the 5'-end sequence/3'-end sequence is selected from those defined in  
 CC the specification. The primer sets can be used in antisense therapy and  
 CC in gene therapy. The primers are useful for synthesizing polynucleotides,  
 CC particularly full-length cDNAs. The primers are also useful for the  
 CC detection and/or diagnosis of the abnormality of the proteins encoded by  
 CC the full-length cDNAs. The primers allow obtaining of the full-length  
 CC cDNAs easily without any specialised methods. AAH03166 to AAH13628 and  
 CC AAH13633 to AAH18742 represent human cDNA sequences; AAB82446 to  
 CC AAB85893 represent human amino acid sequences; and AAH13629 to AAH13632  
 CC represent oligonucleotides, all of which are used in the exemplification  
 CC of the present invention.

XX Sequence 765 BP; 160 A; 236 C; 242 G; 123 T; 4 other;

XX Query Match 22.6%; Score 94.4; DB 22; Length 765;

XX Best Local Similarity 55.5%; Pred. No. 7.7e-12;  
 XX Matches 182; Conservative 0; Mismatches 146; Indels 0; Gaps 0;

QY 42 GGACGAGAGACACTGTGGAGATATGAGAGCCACCGCCACAGATCTGATCAT 101  
 DB 161 GAAACATGACGAGTCTGAAAGCTCTGAGGGCTTCCGGGTGAGCTCCTCGTCT 220  
 QY 102 CTGCCACGAGCCGCTCAACCTTACCTGGGCGAGGCGCAAGGCTGTGACAGTGAAGA 161  
 DB 221 GACCCCTTACGACATCACTTACCTGGGCGAGGCGCAAGGCTGTGACAGTGAAGA 280  
 QY 162 GGAGAGAGTGTCTGACAGCCCGGCTCAACAAGCGCCATGCGGGCGGCACTTCT 221  
 DB 281 GGAGAGAGTGTCTGACAGCCCGGCTCAACAAGCGCCATGCGGGCGGCACTTCT 340  
 QY 222 GATTTGCTGAAGACTCGAGGGAAGACGGGGCCATCGCTTCTGAGAGCTGAAGTT 281  
 DB 341 GGACATCTCTGACGAGCGGCGCCACAAAGGGCTACCTGCTTCTCGAGAGCTGAGCT 400  
 QY 282 CCACAACCTGACGCTACACCTGTGACCCGGGCTGACGCTGATGTTGACTTCAAGTAA 341  
 DB 401 CTACTACCCGAGCTGTACAGAAAGTCAACAGCAAGGACCGGCGGCTTCTTCCAT 460  
 QY 342 CTTTACGGGTGAGAGCTCCGACTTGAAC 369  
 DB 461 GATCATCGACGCTCCGGGAGTCAAGC 488

RESULT 12

AA05387  
 ID AA05387 standard; cDNA; 2098 BP.

XX AA05387;

DT 12-SEP-2001 (first entry)

XX Human caspase recruitment domain, CARD-9 cDNA sequence.

XX Human; caspase recruitment domain; CARD-9; Bcl-10; NF-kappaB;

XX apoptosis; hyperproliferative disorder; autoimmune; neurological;  
 XX inflammatory disorder; viral infection; stress-related response; ss.

OS Homo sapiens.

XX Key Location/Qualifiers  
 FT 144..1754  
 FT CDS /\*tag= a  
 FT /product= "CARD-9"  
 FT /note= "ORF is specifically claimed"

XX W0200140468-A2.

XX 07-JUN-2001.

XX 01-DEC-2000; 2000NO-US32716.

XX 03-DEC-1999; 99US-0168780.

XX 18-FEB-2000; 2000US-0507533.

XX 25-FEB-2000; 2000US-0513904.

XX 10-OCT-2000; 2000US-0685791.

XX (MILL-) MILLENNIUM PHARM INC.

XX Bertin J;

XX WPI; 2001-367809/38.

XX P-PSDB; AAU01205.

XX Claim 2; Fig 5A-5B; 145pp; English.

XX The present sequence encoding for novel human caspase recruitment  
 CC domain, CARD-9 is isolated from a human megakaryocyte cDNA library.  
 CC Also described are novel human sequences for CARD-10 and CARD-11  
 CC (AAU01206, AAU01207) and rat CARD-9 (AAU01204). CARD-9, CARD-10 and  
 CC CARD-11 interact with Bcl-10 which is thought to activate nuclear factor  
 CC (NF)-kappaB and apoptosis. The sequences of the invention can be used for  
 CC treating a disorder associated with abnormal levels of apoptosis by  
 CC modulating the expression or activity of CARD-9, CARD-10, or CARD-11.  
 CC They can be used for the treatment of hyperproliferative disorders  
 CC (e.g. cancer), autoimmune disorders (e.g. systemic lupus erythematosus),  
 CC neurological disorders (e.g. Alzheimer's disease), inflammatory disorders  
 CC (e.g. Crohn's disease), and viral infection (e.g. HIV). The CARD  
 CC polypeptide, polynucleotide and an antibody which selectively binds to  
 CC CARD can be used in screening and detection assays (e.g. chromosomal  
 CC mapping, tissue typing), predictive medicine (prognostic assays,  
 CC monitoring clinical trials, and therapy (treatment and prophylaxis). The  
 CC CARD polypeptide may be used to screen for drugs that bind to and/or  
 CC modulate it. CARD sequences are potential targets for regulating  
 CC inflammation, cancer, NF-kappaB signalling, stress-related response and  
 CC apoptosis in human disease. A host cell containing a polynucleotide  
 CC encoding CARD can be used to create transgenic animals.

XX Sequence 2098 BP; 455 A; 644 C; 702 G; 297 T; 0 other;

XX Query Match 22.6%; Score 94.4; DB 22; Length 2098;

XX Best Local Similarity 55.5%; Pred. No. 8.8e-12;  
 XX Matches 182; Conservative 0; Mismatches 146; Indels 0; Gaps 0;

QY 42 GGACGAGAGACACTGTGGAGATATGAGAGCCACCGCCACAGATCTGATCAT 101

DB 158 GAAACATGACGAGTCTGAAAGCTCTGAGGGCTTCCGGGTGAGCTCCTCGTCTAT 217

QY 102 CTGCCACGAGCCGCTCAACCTTACCTGGGCGAGGCGCAAGGCTGTGACAGTGAAGA 161

DB 218 GACCCCTTACGACATCACTTACCTGGGCGAGGCGCAAGGCTGTGACAGTGAAGA 277

QY 162 GGAAGAGTGTCTGACAGCCCGGCTCAACAAGCGCCATGCGGGCGGGAAGTTGCT 221

DB 278 GGAACAGAGTGTCTGACAGCCCGGCTCAACAAGGCTGTGAGTGTGCTCT 337

QY 222 GATTTGCTGAAGACTCGAGGGAAGACGGGGCCATCGCTTCTGAGAGCTGAAGTT 281

DB 338 GGACATCCTGACGGAGCCGACCAAGAGGCTACGTGGCTCTTCTCGAGAGCTTGAGCT 397  
 QY 282 CCACACCCCTGACCTCTACACCCCTGTCACCGGCTGAGCTTATGTTGACTTCACTTA 341  
 DB 398 CTACTACCCCGAGCTGTACAGAAAGTTCACAGGACAGAGCCGCGCTTCTTCAT 457  
 QY 342 CTTAGCGGTGAGAGCTCCGACTTTGAC 369  
 DB 458 GATCATCGACCGCTCCGGGAGTCAAGC 485

## RESULT 13

ABA00332 ID ABA00332 standard; cDNA, 2098 BP.

AC ABA00332;

DT 09-DEC-2002 (first entry)

DE Human CARD-9 cDNA.

XX Gene; rat; human; caspase recruitment domain; CARD-9; CARD-10;  
 KW CARD-11; apoptosis; inflammation; cell growth; cell death;  
 KW lymphocyte activation; cancer; melanoma; autoimmune disease;  
 KW arthritis; neurological disorder; Alzheimer's disease; ss.

XX Homo sapiens.

XX Key Location/Qualifiers

XX CDS 144..1754

XX FT /\*tag= a

XX FT /product= "CARD-9"

XX W0200270652-A2.

XX PD 12-SEP-2002.

XX PF 28-FEB-2002; 2002WO-US06147.

XX PR 02-MAR-2001; 2001US-0798412.

XX PA (MILL-) MILLENNIUM PHARM INC.

XX Bertin J;

XX WPI: 2002-698749/75.

XX P-PSDB: AAG79553.

XX CARD-9, CARD-10 or CARD-11 polypeptides and polynucleotides, useful for  
 PT treating disorders associated with inappropriate apoptosis or  
 PT lymphocyte activation, e.g. cancer -

XX Disclosure; Fig 5; 151pp; English.

XX This sequence encodes human caspase recruitment domain (CARD)-9.  
 CC CARD proteins play roles in apoptotic and inflammatory signalling  
 CC pathways. CARD-9, -10 and -11 participate in the network of  
 CC interactions that modulate caspase activity. They are thought to be  
 CC useful as modulating agents for regulating a variety of cellular  
 CC processes including cell growth and cell death. CARD proteins and  
 CC nucleic acids are useful for treating a disorder associated with  
 CC inappropriate apoptosis or lymphocyte activation or for diagnosing  
 CC subjects having or that are at risk of developing a disorder associated  
 CC with aberrant CARD-9, CARD-10 or CARD-11 expression or activity, such  
 CC as cancer e.g. melanoma, autoimmune disorders e.g. arthritis, or  
 CC neurological disorders e.g. Alzheimer's disease.

XX Sequence 2098 BP, 455 A, 644 C, 702 G, 297 T, 0 other;

XX Query Match 22.6%; Score 94.4; DB 24; Length 2098;

XX Best Local Similarity 55.5%; Pred. No. 8.8e-12;  
 XX Matches 182; Conservative 0; Mismatches 146; Indels 0; Gaps 0;

QY 42 GGACGAGGAGACACTGTGGAGATGAGAGGCCACAGAGATGTACGCTGAT 101  
 DB 158 GAAAGATGACAGTGTGGAACGTCCTGAGGCTTCCGGTGAAGCTTACCTCGAT 217  
 QY 102 CTGCCCCAGCCGCTTACACCCCTTACCTGCGCCAGGCAAGTGTGTGCACTGAGAGA 161  
 DB 218 CAGCCCCCTACCGATACACCTTACTGTGGAGAGGCAAGTGTGTGCAAGTGTGAG 277  
 QY 162 GAGAGAGTGTGACAGAGCCCGGCTACCAAGCCGATGCGGCGCGGACTTGT 221  
 DB 278 GAGAGAGTGTGACAGAGCCCGGCTACCAAGCCGATGCGGCGCGGACTTGT 337  
 QY 222 GATTTGCTGAGACTTCGAGGGAAGACCGGCGCATCGCTTCTGAGAGCTTGAAT 281  
 DB 338 GGACATCTGACCGGACCGGACCAAGGCTACGCTGCGCTTCTGAGAGCTGAGCT 397  
 QY 282 CCACACCCCTGACCTCTACACCCCTGTCACCGGCTGAGCTTATGTTGACTTCACTTA 341  
 DB 398 CTACTACCCCGAGCTGTACAGAAAGTTCACAGGACAGAGCCGCGCTTCTTCAT 457  
 QY 342 CTTAGCGGTGAGAGCTCCGACTTTGAC 369  
 DB 458 GATCATCGACCGCTCCGGGAGTCAAGC 485

## RESULT 14

AAH18321 ID AAH18321 standard; cDNA, 2176 BP.

XX AAH18321;

XX 26-JUN-2001 (first entry)

XX Human cDNA sequence SEQ ID NO:18327.

XX Human; primer; detection; diagnosis; antisense therapy; gene therapy; ss.

XX Homo sapiens.

XX EP1074617-A2.

XX PD 07-FEB-2001.

XX PF 28-JUL-2000; 2000EP-0116126.

XX PR 29-JUL-1999; 99JP-0248036.

XX PR 27-AUG-1999; 99JP-0300253.

XX PR 11-JAN-2000; 2000JP-0118776.

XX PR 02-MAY-2000; 2000JP-0183767.

XX PR 09-JUN-2000; 2000JP-0241899.

XX (HELI-) HELIX RES INST.

XX Oca T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;

XX Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;

XX WPI: 2001-318749/34.

XX Claim 8; SEQ ID 18327; 2537pp + CD ROM; English.

XX The present invention describes primer sets for synthesizing 5602  
 CC full-length cDNAs defined in the specification. Where a primer set  
 CC comprises: (a) an oligo-dT primer and an oligonucleotide complementary  
 CC to the complementary strand of a polynucleotide which comprises one of  
 CC the 5602 nucleotide sequences defined in the specification, where the  
 CC oligonucleotide comprises at least 15 nucleotides; or (b) a combination  
 CC of an oligonucleotide comprising a sequence complementary to the

CC complementary strand of a polynucleotide which comprises a 5'-end  
CC sequence and an oligonucleotide comprising a sequence complementary to a  
CC polynucleotide which comprises a 3'-end sequence, where the  
CC oligonucleotide comprises at least 15 nucleotides and the combination of  
CC the 5'-end sequence/3'-end sequence is selected from those defined in  
CC the specification. The primer sets can be used in antisense therapy and  
CC in gene therapy. The primers are useful for synthesising polynucleotides  
CC particularly full-length cDNAs. The primers are also useful for the  
CC detection and/or diagnosis of the abnormality of the proteins encoded by  
CC the full-length cDNAs. The primers allow obtaining of the full-length  
CC cDNAs easily without any specialised methods. AAH0166 to AAH1628 and  
CC AAH1633 to AAH18742 represent human cDNA sequences; AAB92446 to  
CC AAB95893 represent human amino acid sequences; and AAH13629 to AAH13632  
CC represent oligonucleotides, all of which are used in the exemplification  
CC of the present invention.

Sequence 2176 BP; 464 A; 657 C; 734 G; 321 T; 0 other;

Query Match	22.6%	Score 94.4	DB 22	Length 2176
Best Local Similarity	55.5%	Pred. No. 8.8e-12		
Matches 182; Conservative	0	Mismatches 146	Indels 0	Gaps 0

[illegible]

```
RESULT 15
ABX11432
ID    ABX11432 standard; DNA; 2176 BP.
```

AC ABX11432;

DT 09-MAY-2003 (first entry)

**DNA encoding caspase recruitment domain containing protein, CARD-10X.**

KM CARD; caspase recruitment domain apoptosis; cell adhesion; inflammation;  
KM cytokine receptor signalling; cancer; glioma; carcinoma; adenocarcinoma;  
KM CARD-containing polypeptide associated disorder; sarcoma; melanoma; ds;  
KM hamatoma; leukaemia; lymphoma; keratinocyte hyperplasia; neoplasia;  
KM keloid; benign prostatic hypertrophy; inflammatory hyperplasia; fibrosis;  
KM restenosis; allergy; arthritis; lupus; Sjogren's syndrome; sepsis; gene;  
KM Crohn's disease; ulcerative colitis; graft versus host disease; stroke;  
KM abnormal cell death disease; myocardial infarction; heart failure;  
KM neurodegenerative disease; Parkinson's disease; Alzheimer's disease;  
KM CARD-10; caspase activator; caspase inhibitor.

OS Unidentified.

Key	Location/Qualifiers
FT	147.1247
FT	
FT	/tag= a

FT	/product= "CARD-10X"
FT	180,.395
FT	misc_feature
FT	/*tag= b
FT	/note= "Encodes caspase recruitment domain (CARD).
FT	Specifically claimed in claim 2"
FT	606,.1037
FT	misc_feature
FT	/*tag= c
FT	/note= "Encodes filament domain. Specifically claimed in
FT	claim 2"

PN US2002164703-A1

PD 07-NOV-2002.

PF 19-DEC-2001; 2001US-0032159.

PR 21-DEC-2000; 2000US-257457P.

PA (PAWL/) PAWLOWSKI K.  
PA (PAWL/) PAWLOWSKI K.

PA (GODZ/) GODZIK A.  
VY

PI Pawlowski K, Reece

DR WPI; 2003-288137/28.

XX

PT and treatment of

PT failure and AIDS

PS Disclosure; Fig 4; 34pp; English.

The invention relates to an isolated nucleic acid molecule encoding a caspase recruitment domain (CARD) containing polypeptide, CARD containing polypeptides are involved in apoptosis (as caspase activators and caspase inhibitors), cell adhesion, inflammation and cytokine receptor signalling. The methods and compositions of the present invention are useful for the diagnosis and treatment of disorders associated with the aberrant expression or activity of the CARD containing polypeptide such as cancer, glioma, carcinoma, adenocarcinoma, sarcoma, melanoma, keloid, hematoma, leukaemia, lymphoma, keratinocyte hyperplasia, neoplasia, benign prostatic hypertrophy, inflammatory hyperplasia, fibrosis, xeroderma, allergies, inflammatory diseases such as arthritis, lupus, Sjogren's syndrome, Crohn's disease, ulcerative colitis, graft versus host disease, sepsis, abnormal cell death diseases such as stroke, myocardial infarction, heart failure, neurodegenerative diseases like Parkinson's disease and Alzheimer's disease, and HIV infection. The present sequence represents DNA encoding the caspase recruitment domain containing protein, CARD-10X.

**Sequence** 2176 BP; 464 A; 657 C; 734 G; 321 T; 0 other;

Query Match	22.6%;	Score 94.4;	DB 25;	Length 2176;
Best Local Similarity	55.5%;	Pred. No. 8.8e-12;		
Matches 182; Conservative	0;	Mismatches 146;	Indels 0;	Gaps 0

QY	42	GGACGAGGAGACACTGTGGGAGATGATGAAAGCCAGCCACAGATTCGTAGCTGCAT	109
Db	161	GAACGATGACGATGCTGTGAACGTCCTGAAAGGCTCCGGGTGACGCTCACCTCGGTTCAT	222
QY	102	CTGACCCAGACCGCCTCACCCCTTACTGCGCAGGCGCAAGGTGCTGTGTCACGCTGAGACGA	161
Db	221	CGACCCCTGACCGATTCACACTTCTCTGGCGAGTGCAGAGTCTCTGAACCTGATGATGA	283
QY	162	GAGGAGGTGCTGCACAAGCCCCGGCTCACCAAGCGCCATGCGGGCCGGGCACTTGCT	221
Db	281	GGAGCAGGTGCTCAGACGACCCCAACCTGTCATCCGCAAAACGAAAGTGGGTGTGCTCCT	340
QY	222	GGAATTTGCTGAAGATCTCAGAGGAAAGAACGGGGCCATCGCTTCTTGAGAGCTGAAATT	281
Db	341	GGAACATCTTCAGCGGAGCCGCGCAACAAAGGGTCTAGCTGGCTCTTCTTCAGAACCTCGAGCT	400

Qy	282	CCACACCCCTGACGCTTACACACCTGTACCCGGGCTGACGCTGATGTTGACTTCAGTAA	341
Db	401	CTACTACCCCGCAGCTGTTCANAGAAAGTTCACAGGCAAGAGCCGCGCTTCTCCAT	460
Qy	342	CTTAGCGGTGAGAGCTCCGACTTTGAC	369
Db	461	GATCATCGACGCGTCCGGGAGTCAGGC	488

Search completed: February 17, 2004, 21:54:30  
 Job time : 243.091 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 17, 2004, 21:44:36 ; Search time 2341.94 Seconds  
(without alignments)  
4327.596 Million cell updates/sec

Title: US-10-032-159A-15

Perfect score: 417  
Sequence: 1 atg999ggaactgfcgagcag.....ggtctctgtraccacagm 417

Scoring table: IDENTITY NUC  
Gapop 10.0, Gapext 1.0

Searched: 22781392 seqs, 1215238056 residues

Total number of hits satisfying chosen parameters: 45562784

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%  
Listing first 45 summaries

Database :  
1: em\_estba:\*  
2: em\_esthum:\*  
3: em\_estin:\*  
4: em\_estmu:\*  
5: em\_estov:\*  
6: em\_estpl:\*  
7: em\_estro:\*  
8: em\_hci:\*  
9: gb\_est1:\*  
10: gb\_est2:\*  
11: gb\_hci:\*  
12: gb\_est3:\*  
13: gb\_est4:\*  
14: gb\_est5:\*  
15: em\_estfun:\*  
16: em\_estom:\*  
17: em\_gss\_hum:\*  
18: em\_gss\_inv:\*  
19: em\_gss\_pln:\*  
20: em\_gss\_vrt:\*  
21: em\_gss\_fun:\*  
22: em\_gss\_mam:\*  
23: em\_gss\_mus:\*  
24: em\_gss\_pro:\*  
25: em\_gss\_rod:\*  
26: em\_gss\_phg:\*  
27: em\_gss\_vrt1:\*  
28: gb\_gss1:\*  
29: gb\_gss2:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	351	84.2	931	13	BU956261 AGENCOURT
2	278	66.7	549	10	BF079819 MA
3	264.6	63.5	1600	11	AK086176 Mus muscu
4	219.6	52.7	806	10	BF100075

5	153.2	36.7	492	12	BQ031835
6	118.8	28.5	1033	9	AU090784
7	117.8	28.2	920	14	CA489572
8	115	27.6	392	10	BF717357
9	112	26.9	618	14	CA589549
10	111.8	26.8	1141	10	BE867544
11	110.2	26.4	732	14	CD349577
12	110.2	26.4	846	10	BE287025
13	106.6	25.6	545	9	AJ455484
14	106.6	25.6	769	9	AJ453294
15	105.8	25.4	584	9	AJ455489
16	102	24.5	655	12	BF649379
17	101.4	24.3	494	10	BF601607
18	94.4	22.6	722	12	BF905841
19	94.4	22.6	738	12	BF909365
20	94.4	22.6	765	9	AU142752
21	94.4	22.6	775	12	BF907340
22	94.2	22.6	739	13	BF975798
23	92.8	22.3	819	12	BF524042
24	92.8	22.3	829	10	BF307399
25	92.8	22.3	829	10	BF307399
26	92.8	22.3	1104	13	BF358485
27	92.8	22.3	1132	12	BF922630
28	91.8	22.0	455	12	BF821044
29	91.8	22.0	479	12	BF148244
30	91.4	21.9	451	14	CB786793
31	90.4	21.7	755	12	BF1912159
32	80.4	19.3	768	12	BF1837662
33	79	18.9	678	12	BF1648452
34	74	17.7	425	14	CB696535
35	73.2	17.6	677	10	BF474019
36	72.4	17.4	601	10	BE291786
37	70.8	17.0	1024	14	BF702884
38	70.8	17.0	1075	11	AK002346
39	67.8	16.3	945	29	CNS021UH
40	66.8	16.0	438	14	CA389497
41	62.2	14.9	465	29	CNS041UM
42	60.6	14.5	238	29	CNS04M7L
43	59.6	14.3	736	29	CNS037FO
44	59.2	14.2	716	29	CNS02NN7
45	56	13.4	566	10	BE040888

## ALIGNMENTS

RESULT 1  
LOCUS BU956261  
DEFINITION AGENCOURT 10613482 NIH MGC 107 Homo sapiens CDNA clone  
IMAGE:6725809 5', mRNA sequence.

ACCESSION BU956261  
VERSION BU956261.1 GI:24185833  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
REFERENCE NIH-MGC http://mgc.ncbi.nlm.nih.gov/  
AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)  
TITLE Unpublished  
JOURNAL Contact: Robert Strausberg, Ph.D.  
COMMENT Email: cgabs-remail.nih.gov

CDNA Library Preparation: Rubin Laboratory  
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)  
DNA Sequencing by: Agencourt Bioscience Corporation  
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at:  
http://image.llnl.gov  
Plate: L1CM3053 row: j column: 16  
High quality sequence stop: 594.



## FEATURES

Location/Qualifiers  
1. 931

/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:6729809"  
/tissue\_type="adenocarcinoma, cell line"  
/lab\_host="DH10B (phage-resistant)"  
/clone\_lib="NIH\_MGC\_107"  
/note="Organ: Breast; Vector: pOTB7, Site 1: EcoRI;  
Site 2: XhoI; cDNA made by oligo-dT priming.  
Directionally cloned into EcoRI/XhoI sites using the  
following 5' adaptor: GGCACGAG(G). Library constructed by  
Ling Hong in the laboratory of Gerald M. Rubin (University  
of California, Berkeley) using ZAP-cDNA synthesis kit  
(Stratagene) and Superscript II RT (Life Technologies).  
Note: this is a NIH\_MGC library."

BASE COUNT 201 a 294 c 295 g 140 t 1 others  
ORIGIN

Query Match 84.2%; Score 351; DB 13; Length 931;  
Best Local Similarity 100.0%; Pred. No. 2.6e-66;  
Matches 351; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGGGGGAATGTCGCGGAGGACTCCGCACTACGCGCACTGAGAGAGACACTGTGG 60  
DB 184 ATGGGGGAATGTCGCGGAGGACTCCGCACTACGCGCACTGAGAGAGACACTGTGG 243  
QY 61 GAGATGATGAG 120  
DB 244 GAGATGATGAG 303  
QY 121 CCTTACTGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 180  
DB 304 CCTTACTGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 363  
QY 181 CCCCAGCTCACCAACAGCGCATGCGGCGCGGCGCATCTGCTGATTTCTGAAGACTGA 240  
DB 364 CCCCAGCTCACCAACAGCGCATGCGGCGCGGCGCATCTGCTGATTTCTGAAGACTGA 423  
QY 241 GGGAGAGAGCGGGGCGCATGCGCTTCTGAGAGAGCTGAAGTTCCAAACCTTGAAGCTTAC 300  
DB 424 GGGAGAGAGCGGGGCGCATGCGCTTCTGAGAGAGCTGAAGTTCCAAACCTTGAAGCTTAC 483  
QY 301 ACCCTGTGTCACCGGGGCTGACGCTGATGTGACTGACTTAACCTTGAAGCGGT 351  
DB 484 ACCCTGTGTCACCGGGGCTGACGCTGATGTGACTGACTTAACCTTGAAGCGGT 534

RESULT 2 549 bp mRNA linear EST 18-OCT-2000  
LOCUS BF079819  
DEFINITION 230513 MARC 2P1G Sus scrofa cDNA 5', mRNA sequence.  
ACCESSION BF079819  
VERSION BF079819.1 GI:10873649  
KEYWORDS EST.  
SOURCE Sus scrofa (pig)  
ORGANISM Sus scrofa

## REFERENCE

FAHRENKRUG, S.C., SMITH, T.P.L., FREKING, B.A., CHO, J., WHITE, J.,  
VALLE, J., WISE, T., ROITER, G.A., PETTEE, G., SULTANA, R., QUACKENBUSH  
'J. and KEELE, J.W.  
Porcine gene discovery by normalized cDNA-library sequencing and  
EST cluster assembly

## TITLE

Mamm. Genome 13 (8), 475-478 (2002)

## JOURNAL

22213789

## MEDLINE

12226715

## COMMENT

Contact: Smith TPL  
USDA, ARS, US Meat Animal Research Center  
PO Box 166, Clay Center, NE 68933-0166, USA  
Tel: 402 762 4366

## FEATURES

source  
1. 549

Fax: 402 762 4390  
Email: smith@emil.marc.usda.gov  
Single pass sequencing. Bases called and alt trimmed with phred  
v0.980904.e. Vector identified by cross\_match with the -mismore 18  
and -mismatch 12 options.  
PCR primers  
FORWARD: AGGAACAGCATGATGACAT  
BACKWARD: GTTTTCCCATCTCACAGC  
Plate: 47 row: M column: 12  
Seq primer: ATTAGGTGACACTATAG.  
Location/Qualifiers

BASE COUNT 102 a 169 c 188 g 90 t  
ORIGIN

Query Match 66.7%; Score 278; DB 10; Length 549;  
Best Local Similarity 87.1%; Pred. No. 1.6e-50;  
Matches 305; Conservative 0; Mismatches 45; Indels 0; Gaps 0;

QY 1 ATGGGGGAATGTCGCGGAGGACTCCGCACTACGCGCACTGAGAGAGACACTGTGG 60  
DB 71 ATGGGGGAATGTCGCGGAGGACTCCGCACTACGCGCACTGAGAGAGAGACTGTGG 130  
QY 61 GAGATGATGAG 120  
DB 131 GAGATGATGAG 190  
QY 121 CCTTACTGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 180  
DB 191 CCTTACTGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 250  
QY 181 CCCCAGCTCACCAACAGCGCATGCGGCGCGGCGCATCTGCTGATTTCTGAAGACTGA 240  
DB 251 CCCCAGCTCACCAACAGCGCATGCGGCGCGGCGCATCTGCTGATTTCTGAAGACTGA 310  
QY 241 GGGAGAGAGCGGGGCGCATGCGCTTCTGAGAGAGCTGAAGTTCCAAACCTTGAAGCTTAC 300  
DB 311 GGGAGAGAGCGGGGCGCATGCGCTTCTGAGAGAGCTGAAGTTCCAAACCTTGAAGCTTAC 370  
QY 301 ACCCTGTGTCACCGGGGCTGACGCTGATGTGACTGACTTAACCTTGAAGCGGT 350  
DB 371 ACCCTGTGTCACCGGGGCTGACGCTGATGTGACTGACTTAACCTTGAAGCGGT 420

RESULT 3 1600 bp mRNA linear HTC 05-DEC-2002  
LOCUS AK086176  
DEFINITION Mus musculus 15 days embryo head cDNA, RIKEN full-length enriched  
library, clone:DB3001019 product:CDSPASE RECRUITMENT DOMAIN  
PROTEIN 14 (BDL10-INTERACTING MAGUK PROTEIN 2) (BIMF2) [Mus  
musculus], full insert sequence.

ACCESSION AK086176  
VERSION AK086176.1 GI:26103277  
KEYWORDS HTC; CAP trapper.  
SOURCE Mus musculus (house mouse)  
ORGANISM Mus musculus

## REFERENCE

CARNINCI, P. and HAYASHIZAKI, Y.  
High-efficiency full-length cDNA cloning  
Meth. Enzymol. 303, 19-44 (1999)  
JOURNAL 99279253  
MEDLINE 10349636  
PUBMED



Tissue Procurement: Gilbert Smith, Ph.D.  
 CDNA Library Preparation: Life Technologies, Inc.  
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (ILNL)  
 DNA Sequencing by: Incyte Genomics, Inc.  
 Clone distribution: MGC clone distribution information can be  
 found through the I.M.A.G.E. Consortium/ILNL at:  
<http://image.llnl.gov>  
 Plate: LLAM9174 row: j column: 12  
 High quality sequence stop: 668.

## FEATURES

source

Location/Qualifiers

1. .806

/organism="Mus musculus"

/mol\_type="mRNA"

/strain="FVB/N"

/db\_xref="taxon:10090"

/clone="IMAGE:3979883"

/issue\_type="tumor, biopsy sample"

/dev\_stage="10 months, virgin"

/lab\_host="DH10B"

/clone\_lib="NCI CGAP Mam1"

/note="Organ: mammary; Vector: pCMV-Sport6; Site 1: SalI;  
 Site 2: NotI; Cloned unidirectionally. Primer: Oligo dt.  
 Library constructed by Life Technologies. Investigator  
 providing samples: Gilbert Smith, NIH"

BASE COUNT

188 a 214 c 242 g 162 t

ORIGIN

Query Match

52.7%; Score 219.6; DB 10; Length 806;

Best Local Similarity 83.6%; Pred. No. 8.7e-38;

Matches 296; Conservative 0; Mismatches 54; Indels 4; Gaps 4;

QY 1 ATGGGGGAAGTGTGGCGGAGGAGTCCGCACTCAACGCGACTGACGAGAGACACTGTGG 60

DB 304 ATGGGGAAGTGTGGCGGAGGAGTCCGCACTCAACGCGACTGACGAGAGACACTGTGG 363

QY 61 GAGATGATGAGAGACACCGCCACAGATGCTGACGCTGATCTGCCCCGCGCTCACC 120

DB 364 GATATGCTGAGAGTACCGATGACAGATGTCAGAGATCTGCCCTGACCGGCTCACC 423

QY 121 CCTTACCTGGCGGAGGAGGAGTGTGCGACGCTGACGAGAGAGGAGTGTGCGACGCG 180

DB 424 CCTTACCTGGCGGAGGAGGAGTGTGCGACGCTGACGAGAGAGGAGTGTGCGACGCG 483

QY 181 CCCCCGCTACCAACAGCGCCATGCGGCGCGGCGCACTTCTGATTTGCTGAAGATCGA 240

DB 484 TCCCGTTTACCAACAGTGC-CATGAGAGTTGGGCACTTCTGATTTGCTGAAGATCGA 542

QY 241 GGGAGAGACGGGGCCATCGCTTCTGAGAGCTGAGATTCACAA-CCCTGACGCTTA 299

DB 543 GGGAGAGAGTGGAGCCATTCCTTCTGGAAGGCTGAGATTCACAAAGCCCTGATGTCTA 602

QY 300 CACCCCTGCTAC-CGGGCTGACAGCTGA-TGTTGATTTAGTACTTTAGCGGT 351

DB 603 CACCCCTGCTACCGGGCTGACAGCTGAGATTTGACTTACGACCTTTAGCGGT 656

RESULT 5  
 B0031835/c 492 bp mRNA linear EST 27-MAR-2002  
 LOCUS  
 DEFINITION  
 UI-1-CFO-aoh-c-03-0-UI.s1 NCI CGAP P171 Mus musculus cDNA clone  
 B0031835  
 ACCESSION  
 VERSION  
 B0031835.1 GI:19767114  
 KEYWORDS  
 EST.  
 SOURCE  
 Mus musculus (house mouse)  
 ORGANISM  
 Mus musculus  
 BUKARYOZA: Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 1 (bases 1 to 492)  
 NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.  
 National Cancer Institute, Cancer Genome Anatomy Project (CGAP),  
 Tumor Gene Index  
 Unpublished

## COMMENT

Contact: Robert Strausberg, Ph.D.  
 Email: [cgaps-remail.nih.gov](mailto:cgaps-remail.nih.gov)  
 Tissue Procurement: Dr. Minoru Ko, Dr. Janet Rosent, Dr. Janet  
 Rosent  
 CDNA Library Preparation: Dr. M. Bento Soares, University of Iowa  
 CDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa  
 DNA Sequencing by: Dr. M. Bento Soares, University of Iowa  
 Clone distribution: Clone distribution information can be obtained  
 from Dr. M. Bento Soares, [bento-soares@uiowa.edu](mailto:bento-soares@uiowa.edu)  
 The following repetitive elements were found in this CDNA  
 sequence: 1-49, >AT rich#low\_complexity (matched complement)  
 Seq primer: M13 FORWARD  
 POLYA=Yes

## FEATURES

source

Location/Qualifiers

1. .492

/organism="Mus musculus"

/mol\_type="mRNA"

/strain="C57BL/6J"

/db\_xref="taxon:10090"

/clone="UI-1-CFO-aoh-c-03-0-UI"

/issue\_type="trophoblast"

/lab\_host="DH10B (Life Technologies)"

/clone\_lib="NCI CGAP P171"

/note="Organ: placenta; Vector: pT73-Pac (Pharmacia) with  
 a modified polylinker; Site 1: EcoR I; Site 2: Not I;  
 NCI CGAP P171 is a subcloned cDNA library constructed  
 according to Bonaldo, Lennon and Soares, Genome Research,  
 6:791-806, 1996. First strand cDNA synthesis was primed  
 with an oligo-dT primer containing a Not I site. Double  
 stranded cDNA was ligated to an EcoR I adaptor, digested  
 with Not I, and cloned directionally into pT73-Pac  
 vector. The oligonucleotide used to prime the synthesis of  
 first-strand cDNA contains a library tag sequence that is  
 located between the Not I site and the (dT)18 tail. The  
 sequence tags for this library are GGTGG, ATCAT, GGGTG,  
 GGTGG. For additional information, contact: Bento Soares,  
 bento-soares@uiowa.edu  
 TAG LIB=UI-1-CFO  
 TAG TISSUE=placenta mouse 11.5-12.5 days  
 TAG\_SEQ=ATCAT"

BASE COUNT

135 a 113 c 129 g 115 t

ORIGIN

Query Match

36.7%; Score 153.2; DB 12; Length 492;

Best Local Similarity 79.1%; Pred. No. 2e-23;

Matches 182; Conservative 0; Mismatches 48; Indels 0; Gaps 0;

QY 1 ATGGGGGAAGTGTGGCGGAGGAGTCCGCACTCAACGCGACTGACGAGAGACACTGTGG 60

DB 430 ATGGGGAAGTGTGGCGGAGGAGTCCGCACTCAACGCGACTGACGAGAGATCTCTGG 371

QY 61 GAGATGATGAGAGACACCGCCACAGATGCTGACGCTGATCTGCCCCGCGCTCACC 120

DB 370 GATATGCTGAGAGTACCGATGACAGATTTGTGAGAGCATCTGCCCTACCGGCTACC 311

QY 121 CCTTACCTGGCGGAGGAGGAGTGTGCGACGCTGACGAGAGAGGAGTGTGCGACGCG 180

DB 310 CCTTACCTGGCGGAGGAGGAGTGTGCGACGCTGACGAGAGAGGAGTGTGCGACGCG 251

QY 181 CCCCCGCTACCAACAGCGCCATGCGGCGCGGCGCACTTCTGATTTGCT 230

DB 250 TCCCGTTTACCAACAGTGCATAGAGTTGACTCTGTGAGATTTCTTCT 201

RESULT 6  
 A0090784 1033 bp mRNA linear EST 23-JAN-2001  
 LOCUS  
 DEFINITION  
 A0090784 lambda ZAP1-Con A stimulated leukocytes Paralicthys  
 olivaceus cDNA clone JF00180F forward, mRNA sequence.  
 A0090784  
 ACCESSION  
 VERSION  
 A0090784.1 GI:12390824  
 KEYWORDS  
 EST.  
 SOURCE  
 Paralicthys olivaceus (bastard halibut)

## ORGANISM

Paralichthys olivaceus  
Eukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;  
Acanthomorphia; Acanthopterygii; Percomorphia; Pleuronectiformes;  
Pleuronectoidae; Paralichthyidae; Paralichthys.

## REFERENCE

1 (bases 1 to 1033)  
Nam, B., Hirono, T. and Aoki, T.

## AUTHORS

Identification of expressed genes from Japanese flounder  
(Paralichthys olivaceus) leukocytes stimulated with Con A/FMA

## TITLE

## JOURNAL

Unpublished

## COMMENT

Contact: Ikuro Hirono  
Laboratory of Genetics and Biochemistry  
Tokyo University of Fisheries  
Konan 4-5-7, Minato-ku, Tokyo 108, Japan  
Email: hirono@tokyo-u-fish.ac.jp.

## FEATURES

## SOURCE

Location/Qualifiers

1..1033

/organism="Paralichthys olivaceus"

/mol\_type="mRNA"

/db\_xref="taxon:8255"

/clone="JFConal180F"

/cell\_type="leukocytes"

/dev\_stage="adult"

/clone\_lib="lambda ZAPII-Con A stimulated leukocytes"

BASE COUNT 311 a 234 c 299 g 189 t

## ORIGIN

Query Match 28.5%; Score 118.8; DB 9; Length 1033;

Best Local Similarity 63.1%; Pred. No. 7.3e-16;

Matches 183; Conservative 0; Mismatches 107; Indels 0; Gaps 0;

QY 34 ACGGCACTGACGAGGAGACACTGGGAGATGATGAGAGCCACCCCAAGATGTA 93

DB 162 AGCTGTATGATGTGAGAGAGCTGTGGAGACGTTGAGAGCAACGCTGAACTAGT 221

QY 94 CGCTGCATCTGCCAGCCGCTCAACCCCTACCTGCGCCAGCCAAAGTGTGTGACAG 153

DB 222 TGCTTCATCTCCCGACAAATCTACCCCTCACTCCGCGAGTCAAAAGTCTGAGTAG 281

QY 154 CTGACGAGGAGAGAGTGTGTCACAGCCCGGCTCAACAGCGCCATGCGGGCGGG 213

DB 282 CAGATGAGGAGAGATTTCTCACTCCCTGCTGCTGCTGCTCAAAAGCAACGAGC 341

QY 214 CACTTGTCTGAGATTGTCTGAGAGCTCGAGGAGAAAGGAGCCATCGCTTCTGAGAGC 273

DB 342 CGCTGCTGCACTCTTCACACTAAAGGAGGAGAGGTTACATAGGCTTTCTGAGAGT 401

QY 274 CTGAAGTTCACAAACCTGACGCTGACACCTGATGACCCGGGCTGACGCC 323

DB 402 CTGAGATTCTACTCCCTGAGCTGTACAAAGCTGTACAGGAGAGATCC 451

## RESULT 7

## CA489572

LOCUS CA489572 920 bp mRNA linear EST 14-NOV-2002

DEFINITION AGENCOURT\_10810481 MAPct. Homo sapiens cDNA clone IMAGE:6722049 5',

## ACCESSION

## CA489572

VERSION CA489572.1 GI:24952363

## KEYWORDS

## EST

## SOURCE

## ORGANISM

## Homo sapiens

## Homo sapiens

## Homo sapiens

## Homo sapiens

## Homo sapiens

## Homo sapiens

## Homo sapiens

## Homo sapiens

## Homo sapiens

## Homo sapiens

## FEATURES

## SOURCE

Location/Qualifiers

1..920

/organism="Homo sapiens"

/mol\_type="mRNA"

/db\_xref="taxon:9606"

/clone="IMAGE:6722049"

/cell\_line="ZR-75-1, MCF7, SK-BR-3, MDA-MB-231, HTERT-HME1

/lab\_host="EMD110B"

/clone\_lib="MAPct"

/note="Vector: pCMV-SPORT6; Site\_1: EcoRV; Site\_2: Not I;

Subcloned with brain, liver, lung, kidney and muscle.

Directionally cloned. Priming method: oligo-dT. Average

insert size: 1800 bp. Library amplification: 26,000 fold.

Kristi A. Egland, James J. Vincent, Robert Strausberg,

Bungkook Lee & Ira Pastan: Discovery of new breast

cancer genes encoding membrane and secreted proteins.

Manuscript submitted."

BASE COUNT 186 a 265 c 345 g 124 t

## ORIGIN

Query Match 28.2%; Score 117.8; DB 14; Length 920;

Best Local Similarity 61.0%; Pred. No. 1.2e-15;

Matches 191; Conservative 0; Mismatches 122; Indels 0; Gaps 0;

QY 42 GAGGAGGAGACACTGTGGAGATGATGAGAGCCACCCGACAGATGCTGCTGAT 101

DB 6 GAGGAGGAGACGCTGTGGAGAGATGAGAGGCGCTCCGATCGCTGCGGCGCT 65

QY 102 CTGCCCCAGCCGCTGACACCCCTACCTGCGCAGGCGCAAGTGTGTGCTGAGACGA 161

DB 66 GAACCCGCGCAAGCTACACCCCTATTTGCGCAAGTCCGGGTATCGACAGAGAGCA 125

QY 162 GAGAGAGTGTGTCACAGCCCGGCTCAACAGCGCCATGCGGGCGGCACTTGTCT 221

DB 126 GAGAGAGTGTGTCAGACCTACCGCTTCCGCGCGCGCTCAACCGGCGGCGCTGAT 185

QY 222 GGAATTGCTGAGAGCTCGAGGAGAAAGGAGCAATGCGCTTCTGAGAGCTGAAGT 281

DB 186 GGAATCTTGGCGTGGCGTGGCAAGAGGGCTATGAGGCTTCTGAGAGCCCTGAGAGT 245

QY 282 CCAACACCTGAGCTTACACCTGCTACCGGGCTGCGAGCCTGATGATTTAGTTA 341

DB 246 CTACTCTCCCGCAACCTTACCGCTGCTACGGGCGCAGAAACCGCCAGCGCTGCCAT 305

QY 342 CTTTACGGGTGAG 354

DB 306 GATCTCTGATGAG 318

## RESULT 8

## BG717357

LOCUS BG717357 392 bp mRNA linear EST 08-MAY-2001

DEFINITION 602689691P1 NIH\_MGC\_97 Homo sapiens cDNA clone IMAGE:4821911 5',

## ACCESSION

## BG717357

VERSION BG717357.1 GI:13996544

## KEYWORDS

## EST

## SOURCE

## ORGANISM

## Homo sapiens

## Homo sapiens

## Homo sapiens

## Homo sapiens

## Homo sapiens

## Homo sapiens

## Homo sapiens

## Homo sapiens

## Homo sapiens

## Homo sapiens

DNA Sequencing by: Agencourt Bioscience Corporation

Clone distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:

http://image.llnl.gov

Plate: LLNL4284 row: c column: 09

High quality sequence stop: 637.

Location/Qualifiers

1..920

/organism="Homo sapiens"

/mol\_type="mRNA"

/db\_xref="taxon:9606"

/clone="IMAGE:6722049"

/cell\_line="ZR-75-1, MCF7, SK-BR-3, MDA-MB-231, HTERT-HME1

/lab\_host="EMD110B"

/clone\_lib="MAPct"

/note="Vector: pCMV-SPORT6; Site\_1: EcoRV; Site\_2: Not I;

Subcloned with brain, liver, lung, kidney and muscle.

Directionally cloned. Priming method: oligo-dT. Average

insert size: 1800 bp. Library amplification: 26,000 fold.

Kristi A. Egland, James J. Vincent, Robert Strausberg,

Bungkook Lee & Ira Pastan: Discovery of new breast

cancer genes encoding membrane and secreted proteins.

Manuscript submitted."

Manuscript submitted."

Manuscript submitted."

Manuscript submitted."

Manuscript submitted."

Manuscript submitted."

Manuscript submitted."

Manuscript submitted."

Manuscript submitted."

Manuscript submitted."

Manuscript submitted."

Manuscript submitted."

Manuscript submitted."

Manuscript submitted."

Manuscript submitted."

Manuscript submitted."

Manuscript submitted."

Manuscript submitted."

Manuscript submitted."

Manuscript submitted."

Manuscript submitted."

Manuscript submitted."

Manuscript submitted."

Manuscript submitted."

Manuscript submitted."

Manuscript submitted."

Manuscript submitted."

Manuscript submitted."

Manuscript submitted."

Manuscript submitted."

Manuscript submitted."

Manuscript submitted."

Manuscript submitted."

Manuscript submitted."

Manuscript submitted."

Manuscript submitted."

Manuscript submitted."

Manuscript submitted."

Manuscript submitted."

Manuscript submitted."

Manuscript submitted."

Manuscript submitted."

Manuscript submitted."

Manuscript submitted."

Manuscript submitted."



BE867544  
LOCUS 1141 bp mRNA linear EST 20-OCT-2000  
DEFINITION 601443043P1 NIH\_MGC\_65 Homo sapiens cDNA clone IMAGE:3847282 5',  
RNA sequence.  
ACCESSION BE867544  
VERSION BE867544.1 GI:10316320  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
1 (bases 1 to 1141)  
NIH-MGC <http://mgs.nci.nih.gov/>.  
National Institutes of Health, Mammalian Gene Collection (MGC)  
JOURNAL Unpublished  
COMMENT Contact: Robert Strausberg, Ph.D.  
Email: [cgabs-remail.nih.gov](mailto:cgabs-remail.nih.gov)  
Tissue Procurement: ATCC  
CDNA Library Preparation: Life Technologies, Inc.  
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNLN)  
DNA Sequencing by: Incyte Genomics, Inc.  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LNLN at:  
<http://image.llnl.gov>  
Plate: L1AM9561 row: e column: 11  
High quality sequence start: 4  
High quality sequence stop: 269.  
Location/Qualifiers  
1..1141  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:3847282"  
/tissue\_type="adenocarcinoma"  
/lab\_host="DH10B (phage-resistant)"  
/clone\_1ib="NIH MGC 65"  
/note="Organ: colon; Vector: pCMV-SPORT6; Site\_1: NotI;  
Site\_2: SalI; Cloned unidirectionally. Primer: Oligo dt.  
Average insert size 1.8 Kb. Library constructed by Life  
Technologies."

BASE COUNT 318 a 351 c 320 g 152 t  
ORIGIN

Query Match 26.8%; Score 111.8; DB 10; Length 1141;  
Best Local Similarity 94.3%; Pred. No. 2.5e-14;  
Matches 116; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 1 ATGGGGGAACCTGCGCGAGGAGCTCCGCACTCCGCACTGAGAGAGAGACACTGTGG 60  
DB 161 ATGGGGGAACCTGTGCGGAGGAGCTCCGCACTCCGCACTGAGAGAGAGACACTGTGG 220  
QY 61 GAGATGATGAGAGAGCCGCGCAAGATCGTACGCTGATGCGCCAGCGGCTCAAC 120  
DB 221 GAGATGATGAGAGAGCCGCGCAAGATCGTACGCTGATGCGCCAGCGCAATTAA 280  
QY 121 CCC 123  
DB 281 CCC 283

RESULT 11  
LOCUS 732 bp mRNA linear EST 29-MAY-2003  
CD349577  
DEFINITION UT-M-FY0-cfr-k-23-0-UI\_r1 NIH\_MGC\_FY0 Mus musculus cDNA clone  
IMAGE: 6855024 5', mRNA sequence.  
ACCESSION CD349577  
VERSION CD349577.1 GI:31141092  
KEYWORDS EST.  
SOURCE Mus musculus (house mouse)  
ORGANISM Mus musculus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
1 (bases 1 to 732)

AUTHORS NIH-MGC <http://mgs.nci.nih.gov/>.  
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
JOURNAL Unpublished  
COMMENT Contact: Robert Strausberg, Ph.D.  
Email: [cgabs-remail.nih.gov](mailto:cgabs-remail.nih.gov)  
Tissue Procurement: Dr. Jim Lin, University of Iowa  
CDNA Library Preparation: Dr. M. Bento Soares, University of Iowa  
CDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa  
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa  
Clone distribution: Distribution information can be found at  
<http://genome.uiowa.edu/distribution/mouse1.html>  
This clone was contributed by the Brain Molecular Anatomy Project  
(BMAP)  
Seq primer: pYX-5.  
Location/Qualifiers  
1..732  
/organism="Mus musculus"  
/mol\_type="mRNA"  
/strain="C57BL/6"  
/db\_xref="taxon:10090"  
/clone="IMAGE: 6855024"  
/tissue\_type="whole brain"  
/dev\_stage="embryo 13.5,14.5,16.5,17.5dpc"  
/lab\_host="DH10B (T1 phage resistant)"  
/clone\_1ib="NIH BMAP FY0"  
/note="Organ: Brain; Vector: pYX-Asc; Site\_1: EcoR I;  
Site\_2: Not I; The library was constructed according  
Bonaldo, Lennon and Soares, Genome Research, 6:791-806,  
1996. Denatured RNA was size fractionated on a 1% agarose  
gel. First strand cDNA synthesis was primed with oligo-dT  
primer containing a Not I site. Double strand cDNA was  
size selected according to mRNA size fraction, ligated  
with EcoR I adaptor, digested with NotI and then cloned  
directionally into pYX-Asc vector. The library tag  
sequence located between the Not I site and the polyA tail  
is AGCGAGACAG. This library was created for the University  
Iowa Brain Anatomy Project (BMAP): 'Gene Discovery in the  
Developing Mouse Nervous System', supported by National  
Institute of Mental Health (NIMH), Hemin Chin, Ph.D.,  
program coordinator."

BASE COUNT 130 a 228 c 277 g 97 t  
ORIGIN

Query Match 26.4%; Score 110.2; DB 14; Length 732;  
Best Local Similarity 61.8%; Pred. No. 4.9e-14;  
Matches 175; Conservative 0; Mismatches 108; Indels 0; Gaps 0;

QY 43 GAGAGAGAGCACTGTGGGAGATGAGAGAGCAAGGCAAGATCGTACGCTGCTATC 102  
DB 309 GAGAGAGAGCAAGCTGTGGGAGAGGAGGAGGCTCCGAGACGATGACGCGGCTTG 368  
QY 103 TGCCCGAGCGGCTCAAGCCCTCACTGCGCCAGGCGCAAGTGTGTGCCAGCTGAGAG 162  
DB 369 AATCCGAGCAAGTCAAGCGGCTATCTGCGCACTGCGGCTCTGAGCAAGAGAGAG 428  
QY 163 GAGAGAGTGTGCAAGAGCCCGGCTCAACCAAGGCGCATGCGGAGCGGAGCACTTGCTG 222  
DB 429 GAGAGAGTGTGCAAGAGCCCGGCTCAACCAAGGCGCATGCGGAGCGGAGCACTTGCTG 488  
QY 223 GATTGTGAAGACTGAGAGAGAGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 282  
DB 489 GACATTTTACGCTGCGGAGGAGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 548  
QY 283 CACAAACCGTGAAGTCAAGCCCTGATCAAGGAGGAGGAGGAGGAGGAGGAGGAGG 325  
DB 549 TACTTACCAGAAACATTCAAGCTGCTACTGAGGAGGAGGAGGAGGAGGAGGAGG 591

RESULT 12  
LOCUS 846 bp mRNA linear EST 26-OCT-2000  
BE287025  
DEFINITION 601097271P1 NCI\_CGAP\_Mam5 Mus musculus cDNA clone IMAGE:3496121 5',  
mRNA sequence.

ACCESSION BE287025  
 VERSION BE287025.1 GI:9165749  
 KEYWORDS EST  
 SOURCE Mus musculus (house mouse)  
 ORGANISM Mus musculus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 REFERENCE NIH-MGC <http://mgc.nci.nih.gov/>  
 1 (bases 1 to 846)  
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
 JOURNAL Unpublished  
 COMMENT Contact: Robert Straubeberg, Ph.D.  
 Email: [cgabbs-remail.nih.gov](mailto:cgabbs-remail.nih.gov)  
 Tissue Procurement: Lothar Hennighausen Ph.D., Robin Humphreys  
 CDNA Library Preparation: Life Technologies, Inc.  
 DNA Sequencing by: Incyte Genomics, Inc.  
 Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/ILNLT at: <http://image.llnl.gov>  
 Plate: LLAM8547 row: m column: 18  
 High quality sequence stop: 697.  
 Location/Qualifiers  
 1. 846  
 /organism="Mus musculus"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:10090"  
 /clone="IMAGE:3496121"  
 /tissue\_type="tumor, gross tissue"  
 /dev\_stage="7 months"  
 /lab\_host="DH10B"  
 /clone\_1lb="NCI\_CGAP\_Mams"  
 /note="Organ: mammary; Vector: pCMV-SPORT6; Site: 1; Salt; Site: 2; Note: Cloned unidirectionally. Primer: Oligo dt. Library constructed by Life Technologies. Investigators providing samples: Lothar Hennighausen/Robin Humphreys, NIH"  
 BASE COUNT 159 a 244 c 314 g 129 t  
 ORIGIN  
 Query Match 26.4%; Score 110.2; DB 10; Length 846;  
 Best Local Similarity 61.8%; Pred. No. 5.1e-14;  
 Matches 175; Conservative 0; Mismatches 108; Indels 0; Gaps 0;  
 QY 43 GACGAGAGACACTGTGGAGATGATGAGAGCCACCCGACAGATCGTACGCTGCATC 102  
 DB 115 GAAAGAGAGCGCTGTGGAGCGATGAGAGGCGTCCGACGACTGACGCGCGCTG 174  
 QY 103 TGCCCCAGCCGCTACCCCTTACTGCGCCAGGCCAAGGTGCTGTGCCAGTGCAG 162  
 DB 175 AATCCGGCCAAAGCTCAGCGCTTACTGCGCCAGTCCGCTCTGACGAAACGAGACGAG 234  
 QY 163 GAGAGAGTGTGCACACCCCGCGCTCACCAAGCGCCATGGGCGCGCGCACTTGCTG 222  
 DB 235 GAGGAGGTGTGCACACCTTACCGTTTCCGTGCGCGCTTACCGTACCGGCGCTTCATA 294  
 QY 223 GATTGCTGAAGACTCGAGGAGAAACGAGGCCATCGCTTCTGTGAGAGCTGAAGTTT 282  
 DB 295 GACATTTTACGCTCCCGGCGAGAGAGGCGCTTGAAGCTTCTTGAAGCCCTTGAAATTC 354  
 QY 283 CACAACCTGACGTCTACACCTGTGTGACCGGCGCTGCAGCTG 325  
 DB 355 TACTACCAAGAACTTCAGCGCTCTCCTGCGCAGGAACTCG 397  
 RESULT 13  
 AJ455484 645 bp mRNA linear EST 22-APR-2002  
 LOCUS AJ455484  
 DEFINITION riken1 Gallus gallus cdna clone 5k17z1, mRNA sequence.  
 ACCESSION AJ455484  
 VERSION AJ455484.1 GI:20265580  
 KEYWORDS EST.

SOURCE Gallus gallus (chicken)  
 ORGANISM Gallus gallus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Gallus.  
 REFERENCE Buerstedde, J.M.  
 1 (bases 1 to 645)  
 TITLE Gallus gallus bursal lymphocyte EST  
 JOURNAL Unpublished  
 COMMENT Contact: Buerstedde JM  
 Cellular Immunology  
 Heinrich-Pette-Institute  
 Martinistr. 52, 20251 Hamburg, Germany  
 Email: <http://genetics.hpi.uni-hamburg.de/dt40est.html>  
 Location/Qualifiers  
 1. 645  
 /organism="Gallus gallus"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9031"  
 /clone="5k17z1"  
 /cell\_type="bursal lymphocyte"  
 /dev\_stage="2-3 weeks old"  
 /clone\_1lb="riken1"  
 /note="CB induced strain"  
 BASE COUNT 166 a 164 c 172 g 142 t 1 others  
 ORIGIN  
 Query Match 25.6%; Score 106.6; DB 9; Length 645;  
 Best Local Similarity 61.2%; Pred. No. 2.9e-13;  
 Matches 172; Conservative 0; Mismatches 109; Indels 0; Gaps 0;  
 QY 43 GACGAGAGACACTGTGGAGATGATGAGAGCCACCCGACAGATCGTACGCTGCATC 102  
 DB 234 GAGGAGAGCTTGTGTGAGAAATGTGAGTGCACACCGGCAATGCTAGCGCTTACATC 353  
 QY 103 TGCCCCAGCCGCTTACCCCTTACTGCGCGCCAGGCCAAGGTGCTGTGCCAGTGCAG 162  
 DB 354 AACCAAGTAAAGTACCCCGTACCTGCGGCAATGCAAGTAAATGATGACCAAGATGAA 413  
 QY 163 GAGAGGTGTGCACAGCCCGCGCTACGCAACGCCCATGCGCGCGGCACTTGCTG 222  
 DB 414 GATGAGGTGTCTTAACTCACTTATGCTGCTCTCCAAATTAACGAGCGCGCGCTGCTG 473  
 QY 223 GATTGCTGAAGACTCGAGGAGAAACGGGCGCATCGCTTCTGAGAGCTGAAGTTT 282  
 DB 474 GACATTTCTTACACCAAGGCGCAGAGGGGCTATGTGTTTCTTGGAGAGCTTGAGATT 533  
 QY 283 CACAACCTGACGTCTACACCTGTGTGACCGGCGCTGCAGCC 323  
 DB 534 TACTACCTGAACCTTACAACTGTGTGACAGGAGAGGCC 574  
 RESULT 14  
 AJ453294 769 bp mRNA linear EST 22-APR-2002  
 LOCUS AJ453294  
 DEFINITION riken1 Gallus gallus cdna clone 33024r1, mRNA sequence.  
 ACCESSION AJ453294  
 VERSION AJ453294.1 GI:20263390  
 KEYWORDS EST.  
 ORGANISM Gallus gallus (chicken)  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Gallus.  
 REFERENCE Buerstedde, J.M.  
 1 (bases 1 to 769)  
 TITLE Gallus gallus bursal lymphocyte EST  
 JOURNAL Unpublished  
 COMMENT Contact: Buerstedde JM  
 Cellular Immunology  
 Heinrich-Pette-Institute  
 Martinistr. 52, 20251 Hamburg, Germany  
 Email: <http://genetics.hpi.uni-hamburg.de/dt40est.html>



FEATURES  
Location/Qualifiers  
1..769  
/organism="Gallus gallus"  
/mol\_type="mRNA"  
/db\_xref="taxon:9031"  
/clone="33024r1"  
/cell\_type="bursal lymphocyte"  
/dev\_stage="2-3 weeks old"  
/clone\_1lb="rikeni"  
/note="CB inbred strain"

BASE COUNT 200 a 193 c 198 g 176 t 2 others  
ORIGIN

Query Match 25.6%; Score 106.6; DB 9; Length 769;  
Best Local Similarity 61.2%; Pred. No. 3e-13;  
Matches 172; Conservative 0; Mismatches 109; Indels 0; Gaps 0;

QY 43 GAGAGAGACACGTGGAGATGATGAGAGCCACAGGATCGTACGCTGATC 102  
DB 315 GAGAGAGAGCTTTGTGGAGATGTCAGTGCACCGGCATGAGCCGTTACATC 374  
QY 103 TGCCCGACCCGCTCACCCCTTACCTGGCCAGGCTGTGCTCCAGCTGACGAG 162  
DB 375 AACCCAGCTAGCTGAGCCCGCTACCTGGGCAATGCAAGTATCGATGACCAAGTGA 434  
QY 163 GAGAGGTGCTGCACAGCCCGGCTCACCAAGCCCATGCGGCGGCACTTGCTG 222  
DB 435 GATGAGGTGCTTAATCACTATGCTGCTGCCCAAAATTAAACGAGCAGGCGGCTGCTG 494  
QY 223 GATTGCTGAAGACTCGAGGAGAAAGCGGGCCATCGCTTCTGAGAGCTTGAATTC 282  
DB 495 GACATTTCTTACACCAAGGCGCAGAGGGCTATGTGTGTTCTTGAGAGCTTGAGATT 554  
QY 283 CACAACCTGAGCTTACACCTGATGTCACCGGCTGACGCC 323  
DB 555 TACTACCTGAACTTACAACTGTGTCAGAGGAAAGGCC 595

RESULT 15  
AJ455489 584 bp mRNA linear EST 22-APR-2002  
LOCUS AJ455489 riken1 Gallus gallus cDNA clone SK23r1, mRNA sequence.  
DEFINITION AJ455489  
ACCESSION AJ455489  
VERSION AJ455489.1 GI:20265585  
KEYWORDS EST.  
SOURCE Gallus gallus (chicken)  
ORGANISM Gallus gallus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Archosauria; Aves; Neognathae; Galliformes; Phasianidae;  
Phasianinae; Gallus.  
1 (bases 1 to 584)  
Buerstedde, J.W.  
Gallus gallus bursal lymphocyte EST  
Unpublished  
Contact: Buerstedde JM  
Cellular Immunology  
Heinrich-Pette-Institute  
Martinistr. 52, 20251 Hamburg, Germany  
Email: URL: <http://genetics.hpi.uni-hamburg.de/dt40est.html>.  
Location/Qualifiers  
1..584  
/organism="Gallus gallus"  
/mol\_type="mRNA"  
/db\_xref="taxon:9031"  
/clone="5K23r1"  
/cell\_type="bursal lymphocyte"  
/dev\_stage="2-3 weeks old"  
/clone\_1lb="rikeni"  
/note="CB inbred strain"

BASE COUNT 151 a 152 c 155 g 122 t 4 others  
ORIGIN

Query Match 25.4%; Score 105.8; DB 9; Length 584;

Best Local Similarity 60.5%; Pred. No. 4.1e-13;  
Matches 170; Conservative 0; Mismatches 111; Indels 0; Gaps 0;

QY 43 GAGAGAGACACTGTGGAGATGATGAGAGCCACCGGCACAGATCGTACGCTGATC 102  
DB 294 GAGAGAGAGCTTTGTGGAGATGTCAGTGCACCGGCATGCTGAGCCGTTACATC 353  
QY 103 TGCCCGACCCGCTCACCCCTTACCTGGCCAGGCCAGAGCTGTGCCAGCTGACGAG 162  
DB 354 AACCCAGCTAGCTGAGCCCGCTACCTGGCCAGTGCACCAAGTATCGATGACCAAGTGA 413  
QY 163 GAGAGGTGCTGCACAGCCCGGCTCACCAAGCCCATGCGGCGGCACTTGCTG 222  
DB 414 GATGAGGTGCTTAATCACTTATGCTGCTGCCCAAAATTAAACGAGCAGGCGGNTGCTG 473  
QY 223 GATTGCTGAAGACTCGAGGAGAAAGCGGGCCATGCGCTTCTGAGAGCTTGAATTC 282  
DB 474 GACATTTTMAACCAAGGCCACAGGGCTATGTGANTTTCTTGAGAGCTTGAGATT 533  
QY 283 CACAACCTGAGCTTACACCTGATGTCACCGGCTGACGCC 323  
DB 534 TACTACCTGAACTTACAACTGTGTCAGAGGAAAGGCC 574

Search completed: February 17, 2004, 23:57:11  
Job time : 2349.94 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 17, 2004, 21:45:41 ; Search time 57.1645 Seconds

(without alignments)  
3219.774 Million cell updates/sec

Title: US-10-032-159a-15

Perfect score: 417

Sequence: 1 atggggggaactgtgcgcgag.....ggctctctgttaaccaccagm 417

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 569978 seqs, 220691566 residues

Total number of hits satisfying chosen parameters: 1139956

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued Patents NA.\*  
1: /cgm2\_6/ptodata/2/ina/5A.COMB.seq.\*  
2: /cgm2\_6/ptodata/2/ina/5B.COMB.seq.\*  
3: /cgm2\_6/ptodata/2/ina/6A.COMB.seq.\*  
4: /cgm2\_6/ptodata/2/ina/6B.COMB.seq.\*  
5: /cgm2\_6/ptodata/2/ina/PTUS.COMB.seq.\*  
6: /cgm2\_6/ptodata/2/ina/backfile1.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	45.2	10.8	2043	US-09-252-991A-8759	Sequence 8759, Ap
2	43.6	10.5	2430	US-09-252-991A-12709	Sequence 12709, A
3	43.6	10.5	2607	US-09-252-991A-13249	Sequence 13249, A
4	43.4	10.4	1771	US-08-533-669A-7	Sequence 7, Appl1
5	43.4	10.4	1771	US-08-511-872-1	Sequence 1, Appl1
6	43.4	10.4	1771	US-08-183-861-7	Sequence 7, Appl1
7	43.4	10.4	1771	US-09-022-765-7	Sequence 7, Appl1
8	43.4	10.4	1771	US-09-551-974A-7	Sequence 7, Appl1
9	43.2	10.4	1179	US-09-252-991A-9030	Sequence 9030, Ap
10	42.8	10.3	936	US-09-252-991A-332	Sequence 332, App
11	42.8	10.3	1959	US-09-252-991A-303	Sequence 303, App
12	42.8	10.3	2016	US-09-252-991A-351	Sequence 351, App
13	42.6	10.2	534	US-09-252-991A-5873	Sequence 5873, Ap
14	42.6	10.2	1968	US-09-252-991A-5829	Sequence 5829, Ap
15	42.6	10.2	2190	US-09-252-991A-5779	Sequence 5779, Ap
16	41.8	10.0	2277	US-08-676-967-5	Sequence 5, Appl1
17	41.8	10.0	2277	US-08-676-974-5	Sequence 5, Appl1
18	41.8	10.0	2277	US-09-098-487-5	Sequence 5, Appl1
19	41.2	9.9	987	US-09-252-991A-15247	Sequence 15247, A
20	41.2	9.9	1344	US-09-252-991A-15551	Sequence 15551, A
21	40.8	9.8	1470	US-09-252-991A-2857	Sequence 2857, Ap
22	40.8	9.8	3027	US-09-252-991A-2758	Sequence 2758, Ap
23	40.6	9.7	708	US-09-266-965-39	Sequence 39, Appl1
24	40.6	9.7	4749	US-09-614-034-189	Sequence 189, Appl1
25	40.6	9.7	5300	US-09-266-965-76	Sequence 76, Appl1
26	40.4	9.7	1413	US-09-252-991A-6071	Sequence 6071, Ap
27	40.4	9.7	1491	US-09-252-991A-6232	Sequence 6232, Ap

28	40.4	9.7	1506	US-09-252-991A-6228	Sequence 6228, Ap
29	40.4	9.7	3111	US-09-252-991A-1778	Sequence 1778, Ap
30	40	9.6	3468	US-07-951-715A-2	Sequence 2, Appl1
31	40	9.6	3468	US-08-459-448A-2	Sequence 2, Appl1
32	40	9.6	3468	US-08-459-595A-2	Sequence 2, Appl1
33	40	9.6	3468	US-08-459-504B-2	Sequence 2, Appl1
34	40	9.6	3468	US-08-459-444-2	Sequence 2, Appl1
35	40	9.6	3468	US-09-053-549-3	Sequence 3, Appl1
36	40	9.6	3468	US-09-547-422-2	Sequence 2, Appl1
37	39.8	9.5	786	US-09-252-991A-10793	Sequence 10793, A
38	39.8	9.5	1080	US-09-252-991A-15937	Sequence 15937, A
39	39.8	9.5	2535	US-09-252-991A-16113	Sequence 16113, A
40	39.8	9.5	2721	US-09-252-991A-16144	Sequence 16144, A
41	39.8	9.5	4665	US-09-252-991A-10243	Sequence 10243, A
42	39.6	9.5	492	US-09-252-991A-10064	Sequence 10064, A
43	39.6	9.5	1029	US-09-252-991A-9770	Sequence 9770, Ap
44	39.6	9.5	1182	US-09-252-991A-10561	Sequence 10561, A
45	39.6	9.5			

## ALIGNMENTS

RESULT 1  
US-09-252-991A-8759  
; Sequence 8759, Application US/09252991A  
; Patent No. 6551795  
; GENERAL INFORMATION:  
; APPLICANT: Marc J. Rubenfield et al.  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS  
; FILE REFERENCE: 107196.136  
; CURRENT APPLICATION NUMBER: US/09/252,991A  
; PRIOR FILING DATE: 1999-02-18  
; PRIOR APPLICATION NUMBER: US 60/074,788  
; PRIOR FILING DATE: 1998-02-18  
; PRIOR APPLICATION NUMBER: US 60/094,190  
; PRIOR FILING DATE: 1998-07-27  
; NUMBER OF SEQ ID NOS: 33142  
; SEQ ID NO 8759  
; LENGTH: 2043  
; TYPE: DNA  
; ORGANISM: Pseudomonas aeruginosa  
US-09-252-991A-8759

Query Match 10.8%; Score 45.2; DB 4; Length 2043;  
Best Local Similarity 48.8%; Pred. No. 0.048;  
Matches 122; Conservative 0; Mismatches 128; Indels 0; Gaps 0;

QY	71	AGAGCCACCGCCACAGATCGTACGCTGATCTGCCCGCCGCTACCCCTTACTGC	130
DB	164	ACAGCCCGCCAGCCCTTACGCTGACGCGGCAATGCCACCATCTCACCGCCACCTTC	223
QY	131	GCGAGCCCAAGTGTCTGTCGCTGACGAGAGAGTGTGACAGCCCCCGGCTCA	190
DB	224	ACAACACCCGCCACCAAGCTGCTGACGCGCTGCGGAGATCCGCCAATCCCTCCG	283
QY	191	CCAAACGCCCGCCATGGCGGCGCGGCACTTGCTGTAAGACTCAGGGAAGAAC	250
DB	284	GCGATGCCAGCCCGCCAGCGCTTACAGAGTCTTGGCCCTCTGCTGTGCTGACCC	343
QY	251	GGGCGCATCGCTTCTGTAAGACCTGAAGTTCCACAACTTCACTCAACCTTGTC	310
DB	344	TGGCCGCAACAACCGGGGTTCAGAACCTGCGCACGACGACATCGTCAACCGATCT	403
QY	311	CCGGGCTGCA 320	
DB	404	TGACCCAGCA 413	

RESULT 2  
US-09-252-991A-12709  
; Sequence 12709, Application US/09252991A

Patent No. 6551795  
GENERAL INFORMATION:  
APPLICANT: Marc J. Rubenfield et al.  
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS  
TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS  
FILE REFERENCE: 107196.136  
CURRENT APPLICATION NUMBER: US/09/252,991A  
CURRENT FILING DATE: 1999-02-18  
PRIOR APPLICATION NUMBER: US 60/074,788  
PRIOR FILING DATE: 1998-02-18  
PRIOR APPLICATION NUMBER: US 60/094,190  
PRIOR FILING DATE: 1998-07-27  
NUMBER OF SEQ ID NOS: 33142  
SEQ ID NO 12709  
LENGTH: 2430  
TYPE: DNA  
ORGANISM: Pseudomonas aeruginosa  
FEATURE:  
NAME/KEY: unsure  
LOCATION: (91)  
OTHER INFORMATION: Identity of nucleotide at the above locations are unknown.  
US-09-252-991A-12709

Query Match 10.5%; Score 43.6; DB 4; Length 2430;

Best Local Similarity 48.8%; Pred. No. 0.12;

Matches 118; Conservative 0; Mismatches 124; Indels 0; Gaps 0;

QY 80 GCCACGAGATCGTACGCTGATCTGCGCCGCTCACTCCCTACCTGCGCCAGGCCA 139  
DB 1193 GCAAGACCAAGATCGGCAATCATCGCCGAGAGCTCGCGCGCTTCTACCGTTTCA 1252  
QY 140 AGGTGCTGTGCTGACGCTGAGACGAGAGAGAGTGTGTACAGCCCGGCTCACCACAGCG 199  
DB 1253 GCGTCGCGCGCAATCGCGCAGACGAGCGCGAGATCAAGGCGCACCGCGCCTTACATCGCGG 1312  
QY 200 CCATGCGGCGCGCGGCACTTGCTGATTTGCTGAAGCTCGAGGGAAGAACGGGCGCATCG 259  
DB 1313 CCTGCGCGCGCAAGCTGTGTGACGAGCAATTAAGAGAGTCAAGTGAATACCGCGTGTATCA 1372  
QY 260 CCTTCTGAGAGCTGAGATTTCACAACTGACGCTTACACCTGTACCGCGGCTGCG 319  
DB 1373 TGCTGACGAGATCGACAAAGCTCGCGCGCACTACCAAGGCGACCGCGCTCGCGGCTGCG 1432  
QY 320 AG 321  
DB 1433 TG 1434

RESULT 3  
US-09-252-991A-13249/c  
Sequence 13249, Application US/09252991A  
Patent No. 6551795  
GENERAL INFORMATION:  
APPLICANT: Marc J. Rubenfield et al.  
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS  
TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS  
FILE REFERENCE: 107196.136  
CURRENT APPLICATION NUMBER: US/09/252,991A  
CURRENT FILING DATE: 1999-02-18  
PRIOR APPLICATION NUMBER: US 60/074,788  
PRIOR FILING DATE: 1998-02-18  
PRIOR APPLICATION NUMBER: US 60/094,190  
PRIOR FILING DATE: 1998-07-27  
NUMBER OF SEQ ID NOS: 33142  
SEQ ID NO 13249  
LENGTH: 2607  
TYPE: DNA  
ORGANISM: Pseudomonas aeruginosa  
FEATURE:  
NAME/KEY: unsure  
LOCATION: (2532)  
OTHER INFORMATION: Identity of nucleotide at the above locations are unknown.  
US-09-252-991A-13249

Query Match 10.5%; Score 43.6; DB 4; Length 2607;  
Best Local Similarity 48.8%; Pred. No. 0.12;  
Matches 118; Conservative 0; Mismatches 124; Indels 0; Gaps 0;

QY 80 GCCACGAGATCGTACGCTGATCTGCGCCGCTCACTCCCTACCTGCGCCAGGCCA 139  
DB 1430 GCAAGACCAAGATCGGCAATCATCGCCGAGAGCTCGCGCGCTTCTACCGTTTCA 1371  
QY 140 AGGTGCTGTGCTGACGCTGAGACGAGAGAGTGTGTACAGCCCGGCTCACCACAGCG 199  
DB 1370 GCGTCGCGCGCAATCGCGCAGACGAGCGCGAGATCAAGGCGCACCGCGCCTTACATCGCGG 1311  
QY 200 CCATGCGGCGCGGCACTTGCTGATTTGCTGAAGCTCGAGGGAAGAACGGGCGCATCG 259  
DB 1310 CCTGCGCGCGCAAGCTGTGTGACGAGCAATTAAGAGAGTCAAGTGAATACCGCGTGTATCA 1251  
QY 260 CCTTCTGAGAGCTGAGATTTCACAACTGACGCTTACACCTGTGACCGCGGCTGCG 319  
DB 1250 TGCTGACGAGATCGACAAAGCTCGCGCGCACTACCAAGGCGACCGCGCTCGCGGCTGCG 1191  
QY 320 AG 321  
DB 1190 TG 1189

#### RESULT 4

US-08-533-669A-7/c

Sequence 7, Application US/08533669A

Patent No. 5834592

GENERAL INFORMATION:

APPLICANT: Corixa Corporation

TITLE OF INVENTION: LEISHMANIA ANTIGENS FOR USE IN THE

TITLE OF INVENTION: THERAPY AND DIAGNOSIS OF LEISHMANIASIS

NUMBER OF SEQUENCES: 18

CORRESPONDENCE ADDRESS:

STREET: 6300 Columbia Center, 701 Fifth Avenue

CITY: Seattle

STATE: Washington

COUNTRY: USA

ZIP: 98104-7092

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/533,669A

FILING DATE: 22-SEP-1995

CLASSIFICATION: 424

ATTORNEY/AGENT INFORMATION:

NAME: Markl, David J.

REGISTRATION NUMBER: 31,392

REFERENCE/DOCKET NUMBER: 210121.420

TELECOMMUNICATION INFORMATION:

TELEPHONE: (206) 622-4900

TELEFAX: (206) 682-6031

INFORMATION FOR SEQ ID NO: 7:

SEQUENCE CHARACTERISTICS:

LENGTH: 1771 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

FEATURE:

NAME/KEY: CDS

LOCATION: 1..1698

US-08-533-669A-7

#### Query Match

10.4%; Score 43.4; DB 2; Length 1771;

Best Local Similarity 46.5%; Pred. No. 0.13;

Matches 140; Conservative 0; Mismatches 161; Indels 0; Gaps 0;

QY 19 AGGGAATCCGCACTCAAGCGGAGTGAAGAGAGACATGTGGGAATATGAGAGCCAC 78  
Db 432 AGCGGTCTCTCTCTTCAGAGGAGCCCTGAGAGCGGAGCTGTGAGCGGATGATACATCCAG 373  
QY 79 CGCCACAGAGATGTAAGCTGATCTGTCCCAAGCGGAGCTCAACCCCTACCTGTGGCAGAGCC 138  
Db 372 GCGGAGCGGAGCTCTCGGCTTCTCTCCAGCTCCGCGGAGACCGGCTGCTCTCCAGAGGC 313  
QY 139 AAGGTCTGTGCAAGCTGAG 198  
Db 312 CTGGAGCGGAGCTGTCTCGGAGCTGTGATAGAGCTCAAGCGGAGCGGAGCTCTCGGAGC 253  
QY 199 GCGATCGGAGCGGAGAGCTGTGATTTGCTGAGAGATCTGAGAGAGAGAGAGAGAGAG 258  
Db 252 CTCGAGCTCGGAGCGGAG 193  
QY 259 GCGTTCCTGAGAGAGCTGAAAGTTCACAAACCTGAGAGCTTACACCTGTGACCGGAGCTG 318  
Db 192 CGCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 133  
QY 319 C 319  
Db 132 C 132

## RESULT 5

US-08-511-872-1/C  
Sequence 1, Application US/08511872

Patent No. 5965142

GENERAL INFORMATION:

APPLICANT: Dillon, Davin C.

APPLICANT: Reed, Steven G.

TITLE OF INVENTION: POLYPEPTIDES AND METHODS FOR THE

TITLE OF INVENTION: DETECTION OF L. TROPICA INFECTION

NUMBER OF SEQUENCES: 11

CORRESPONDENCE ADDRESS:

ADDRESSEE: SEED AND BERRY

STREET: 6300 Columbia Center, 701 Fifth Avenue

CITY: Seattle

STATE: Washington

COUNTRY: USA

ZIP: 98104-7092

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent in Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/511,872

FILING DATE: 04-AUG-1995

CLASSIFICATION: 424

ATTORNEY/AGENT INFORMATION:

NAME: No. 5965142tenburg, Carol

REGISTRATION NUMBER: 39,317

REFERENCE/DOCKET NUMBER: 210121.405

TELECOMMUNICATION INFORMATION:

TELEPHONE: (206) 622-4900

TELEFAX: (206) 682-6031

TELEX: 3723836 SEEDANDBERRY

INFORMATION FOR SEQ ID NO: 1:

SEQUENCE CHARACTERISTICS:

LENGTH: 1771 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

FEATURE:

NAME/KEY: CDS

LOCATION: 1..1698

US-08-511-872-1

Query Match 10.4%; Score 43.4; DB 2; Length 1771;

Best Local Similarity 46.5%; Pred. No. 0.13;

Matches 140; Conservative 0; Mismatches 161; Indels 0; Gaps 0;  
QY 19 AGGGAATCCGCACTCAAGCGGAGTGAAGAGAGACATGTGGGAATATGAGAGCCAC 78  
Db 432 AGCGGTCTCTCTCTTCAGAGGAGCCCTGAGAGCGGAGCTGTGAGCGGATGATACATCCAG 373  
QY 79 CGCCACAGAGATGTAAGCTGATCTGTCCCAAGCGGAGCTCAACCCCTACCTGTGGCAGAGCC 138  
Db 372 GCGGAGCGGAGCTCTCGGAGCTGTGATAGAGCTCAAGCGGAGCGGAGCTCTCGGAGC 313  
QY 139 AAGGTCTGTGCAAGCTGAG 198  
Db 312 CTGGAGCGGAGCTGTCTCGGAGCTGTGATAGAGCTCAAGCGGAGCGGAGCTCTCGGAGC 253  
QY 199 GCGATCGGAGCGGAGAGCTGTGATTTGCTGAGAGATCTGAGAGAGAGAGAGAGAGAG 258  
Db 252 CTCGAGCTCGGAGCGGAG 193  
QY 259 GCGTTCCTGAGAGAGCTGAAAGTTCACAAACCTGAGAGCTTACACCTGTGACCGGAGCTG 318  
Db 192 CGCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 133  
QY 319 C 319  
Db 132 C 132

## RESULT 6

US-09-183-861-7/C  
Sequence 7, Application US/09183861

Patent No. 6365165

GENERAL INFORMATION:

APPLICANT: Reed, Steven G.

APPLICANT: Campos-Neto, Antonio

APPLICANT: Webb, John R.

APPLICANT: Dillon, Davin C.

APPLICANT: Skeiky, Yasir A.W.

TITLE OF INVENTION: LEISHMANIA ANTIGENS FOR USE IN THE THERAPY AND

NUMBER OF SEQUENCES: 87

CORRESPONDENCE ADDRESS:

ADDRESSEE: SEED AND BERRY LLP

STREET: 6300 Columbia Center, 701 Fifth Avenue

CITY: Seattle

STATE: Washington

COUNTRY: USA

ZIP: 98104-7092

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent in Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/183,861

FILING DATE:

CLASSIFICATION:

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 09/022,765

FILING DATE: 12-FEB-1998

ATTORNEY/AGENT INFORMATION:

NAME: Maki, David J.

REGISTRATION NUMBER: 31,392

REFERENCE/DOCKET NUMBER: 210121.420C3

TELECOMMUNICATION INFORMATION:

TELEPHONE: (206) 622-4900

TELEFAX: (206) 682-6031

INFORMATION FOR SEQ ID NO: 7:

SEQUENCE CHARACTERISTICS:

LENGTH: 1771 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

FEATURE:

NAME/KEY: CDS

LOCATION: 1..1698  
US-09-183-861-7

Query Match 10.4%; Score 43.4; DB 4; Length 1771;  
Best Local Similarity 46.5%; Pred. No. 0.13;  
Matches 140; Conservative 0; Mismatches 161; Indels 0; Gaps 0;

QY 19 AGGAGCTCCGCACTCAGCGCACTGAGAGAGACACTGTGGAGATGATGAGAGCCAC 78  
DB 432 AGCGCTGCTCTCTCGAGGGCCCTGAGAGCGGGCTCTCGCGCATGATGATCAG 373  
QY 79 CGGCAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 138  
DB 372 GCGGCGCGCTCTCTCGAGGGCCCTGAGAGCGGGCTCTCGCGCATGATGATCAG 313  
QY 139 AGGTGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 198  
DB 312 CTGGAGAGCGGGCTCTCTCGAGGGCCCTGAGAGCGGGCTCTCGCGCATGATGATCAG 253  
QY 199 GCGATGCGGGCGCGGCACTTGTGATTTGCTGAAGACTGAGAGAAAGAGGGCCATC 258  
DB 252 CTCAGCTCGCGCGGAGAGCGGCTCTCTCGAGGGCCCTGAGAGCGGGCTCTCGCG 193  
QY 259 GCGTCTCTGAGAGAGCTGAGATTCACAACTGAGCTGACACCTGAGTCAACGAGGCTG 318  
DB 192 CGCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 133  
QY 319 C 319  
DB 132 C 132

RESULT 7  
US-09-022-765-7/c

Sequence 7, Application US/09022765  
Patent No. 6375955

GENERAL INFORMATION:

APPLICANT: Reed, Steven G.  
APPLICANT: Campos-Neto, Antonio  
APPLICANT: Webb, John R.  
APPLICANT: Dillon, David C.  
APPLICANT: Skeiky, Yasir A.M.  
TITLE OF INVENTION: LEISHMANIA ANTIGENS FOR USE IN THE THERAPY AND  
NUMBER OF SEQUENCES: 87  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: SEED AND BERRY LLP  
STREET: 6300 Columbia Center, 701 Fifth Avenue  
City: Seattle  
STATE: Washington  
COUNTRY: USA  
ZIP: 98104-7092

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent in Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/022,765

FILING DATE: 12-FEB-1998

CLASSIFICATION:

ATTORNEY/AGENT INFORMATION:

NAME: Marki, David J.  
REGISTRATION NUMBER: 31,392  
REFERENCE/DOCKET NUMBER: 210121.42003

TELEPHONE: (206) 622-4900

TELEFAX: (206) 682-6031

INFORMATION FOR SEQ ID NO: 7:

SEQUENCE CHARACTERISTICS:

LENGTH: 1771 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

FEATURE:  
NAME/KEY: CDS  
LOCATION: 1..1698  
US-09-022-765-7

Query Match 10.4%; Score 43.4; DB 4; Length 1771;  
Best Local Similarity 46.5%; Pred. No. 0.13;  
Matches 140; Conservative 0; Mismatches 161; Indels 0; Gaps 0;

QY 19 AGGAGCTCCGCACTCAGCGCACTGAGAGAGACACTGTGGAGATGATGAGAGCCAC 78  
DB 432 AGCGCTGCTCTCTCGAGGGCCCTGAGAGCGGGCTCTCGCGCATGATGATCAG 373  
QY 79 CGGCAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 138  
DB 372 GCGGCGCGCTCTCTCGAGGGCCCTGAGAGCGGGCTCTCGCGCATGATGATCAG 313  
QY 139 AGGTGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 198  
DB 312 CTGGAGAGCGGGCTCTCTCGAGGGCCCTGAGAGCGGGCTCTCGCGCATGATGATCAG 253  
QY 199 GCGATGCGGGCGCGGCACTTGTGATTTGCTGAAGACTGAGAGAAAGAGGGCCATC 258  
DB 252 CTCAGCTCGCGCGGAGAGCGGCTCTCTCGAGGGCCCTGAGAGCGGGCTCTCGCG 193  
QY 259 GCGTCTCTGAGAGAGCTGAGATTCACAACTGAGCTGACACCTGAGTCAACGAGGCTG 318  
DB 192 CGCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 133  
QY 319 C 319  
DB 132 C 132

RESULT 8

US-09-551-974A-7/c

Sequence 7, Application US/09551974A

Patent No. 650437

GENERAL INFORMATION:

APPLICANT: Reed, Steven G.

APPLICANT: Campos-Neto, Antonio

APPLICANT: Webb, John R.

APPLICANT: Dillon, David C.

APPLICANT: Skeiky, Yasir A.M.

TITLE OF INVENTION: LEISHMANIA ANTIGENS FOR USE IN THE

TITLE OF INVENTION: THERAPY AND DIAGNOSIS OF LEISHMANIASIS

FILE REFERENCE: 210121.42005

CURRENT APPLICATION NUMBER: US/09/551,974A

CURRENT FILING DATE: 2000-04-14

NUMBER OF SEQ ID NOS: 101

SOFTWARE: FastSeq for Windows Version 4.0

SEQ ID NO 7

LENGTH: 1771

TYPE: DNA

ORGANISM: Leishmania tropica

US-09-551-974A-7

Query Match 10.4%; Score 43.4; DB 4; Length 1771;  
Best Local Similarity 46.5%; Pred. No. 0.13;  
Matches 140; Conservative 0; Mismatches 161; Indels 0; Gaps 0;

QY 19 AGGAGCTCCGCACTCAGCGCACTGAGAGAGACACTGTGGAGATGATGAGAGCCAC 78  
DB 432 AGCGCTGCTCTCTCGAGGGCCCTGAGAGCGGGCTCTCGCGCATGATGATCAG 373  
QY 79 CGGCAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 138  
DB 372 GCGGCGCGCTCTCTCGAGGGCCCTGAGAGCGGGCTCTCGCGCATGATGATCAG 313  
QY 139 AGGTGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 198  
DB 312 CTGGAGAGCGGGCTCTCTCGAGGGCCCTGAGAGCGGGCTCTCGCGCATGATGATCAG 253

QY 199 GCCATGGGGCCGGGCACTTGTGGATTGCTGTAAGACTCGAGGGAAGAAGCGGCCATC 258  
 Db 252 CTCGACGTCCGGCGGGAGACGCGTGCCTCTCGAGGGCTGACGCGGGCTGTCTGGC 193  
 QY 259 GCCTTCTGGAGAGCTGAAGTTCCACAACCCCTGACGCTTACACCTGTGTCACCGGGCTG 318  
 Db 192 CGCATGATGACATCCAGCGGGCGCGCTCTCGGCTCTTCCAGTCTCCGCGCGAGACG 133  
 QY 319 C 319  
 Db 132 C 132

RESULT 9  
 US-09-252-991A-9030/c  
 ; Sequence 9030, Application US/09252991A  
 ; Patent No. 6551795

GENERAL INFORMATION:

APPLICANT: Marc J. Rubenfield et al.

TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS

TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS

FILE REFERENCE: 107196.136

CURRENT APPLICATION NUMBER: US/09/252,991A

PRIOR FILING DATE: 1999-02-18

PRIOR FILING DATE: 1998-02-18

PRIOR APPLICATION NUMBER: US 60/094,190

PRIOR FILING DATE: 1998-07-27

NUMBER OF SEQ ID NOS: 33142

SEQ ID NO 9030

LENGTH: 1179

TYPE: DNA

ORGANISM: Pseudomonas aeruginosa

US-09-252-991A-9030

Query Match 10.4%; Score 43.2; DB 4; Length 1179;  
 Best Local Similarity 50.0%; Pred. No. 0.13;  
 Matches 108; Conservative 0; Mismatches 108; Indels 0; Gaps 0;

QY 105 CCCGAGCGGCTCAACCCCTAAGTGGCGGCAAGGCTGTGCGAGCTGACGAGGA 164  
 Db 1159 CGCCACCATCTACCGGACCGCTGCAACAAGAGCGGACGAGTCCGCTCGACGGCCT 1100  
 QY 165 GGAAGTGTGACAGCCCGGCTCACCACAGGCGCATGGGGCGGGGACTTGTCTGA 224  
 Db 1099 GGTGCGGAGATCCGCAATCCCGGCGATGCGAGCGCGGATACCAAGGCTGTT 1040  
 QY 225 TTTCGTGAAGCTGAGGGAAGAGAGGCGCATGCGCTTCTGAGAGGCTGAAGTTCCA 284  
 Db 1039 GCGCCCTGCTGCTGTGGTGGCTGACGTGGCCAGCAACCGGGTGTTCAGAACTGGC 980  
 QY 285 CAACCTGAGCTACACCTGTGTCACCGGGCTGCA 320  
 Db 979 CACCAAGACATCTCTACCCAGGTCTTCAACAGCA 944

RESULT 10  
 US-09-252-991A-332/c  
 ; Sequence 332, Application US/09252991A  
 ; Patent No. 6551795  
 GENERAL INFORMATION:  
 APPLICANT: Marc J. Rubenfield et al.  
 TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS  
 TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS  
 FILE REFERENCE: 107196.136  
 CURRENT APPLICATION NUMBER: US/09/252,991A  
 PRIOR FILING DATE: 1999-02-18  
 PRIOR APPLICATION NUMBER: US 60/074,788  
 PRIOR FILING DATE: 1998-02-18  
 PRIOR APPLICATION NUMBER: US 60/094,190  
 PRIOR FILING DATE: 1998-07-27  
 NUMBER OF SEQ ID NOS: 33142  
 SEQ ID NO 332

LENGTH: 936  
 TYPE: DNA  
 ORGANISM: Pseudomonas aeruginosa  
 US-09-252-991A-332

Query Match 10.3%; Score 42.8; DB 4; Length 936;  
 Best Local Similarity 51.6%; Pred. No. 0.15;  
 Matches 98; Conservative 0; Mismatches 92; Indels 0; Gaps 0;

QY 59 GGGAGATGATGAGAGGCAACCGCACAGGATGTAAGCTGATCTGCCACCGGCTCA 118  
 Db 703 GGAAGCGCTGTGATCTGATGATGCGGATGCTCGCGGCTTCAACGCGCGATCG 644  
 QY 119 CCCCTACCTGGCGGAGGCGCAAGTGTCTGCTGACCTGGAACGAGAGAGTGTGACA 178  
 Db 643 ATACCAACCTGGGCGCGGCGGACGACCGCTGCGCGGTGACATGAGCTGCGGTACA 584  
 QY 179 GCGCCCGGCTCAACCAAGCGCATGCGGCGGGCACTTGTGATTTGTGAAGACTC 238  
 Db 583 TCGCCACGTTGGCGCCCACTGAGGTGGCGAGCGCGATGCGGCGGTGAGCAAGCCT 524  
 QY 239 GAGGGAAGA 248  
 Db 523 GTTCGAAGA 514

RESULT 11  
 US-09-252-991A-303  
 ; Sequence 303, Application US/09252991A  
 ; Patent No. 6551795

GENERAL INFORMATION:

APPLICANT: Marc J. Rubenfield et al.

TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS

TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS

FILE REFERENCE: 107196.136

CURRENT APPLICATION NUMBER: US/09/252,991A

PRIOR FILING DATE: 1999-02-18

PRIOR FILING DATE: 1998-02-18

PRIOR APPLICATION NUMBER: US 60/094,190

PRIOR FILING DATE: 1998-07-27

NUMBER OF SEQ ID NOS: 33142

SEQ ID NO 303

LENGTH: 1959

TYPE: DNA

ORGANISM: Pseudomonas aeruginosa

US-09-252-991A-303

Query Match 10.3%; Score 42.8; DB 4; Length 1959;  
 Best Local Similarity 51.6%; Pred. No. 0.18;  
 Matches 98; Conservative 0; Mismatches 92; Indels 0; Gaps 0;

QY 59 GGGAGATGATGAGAGGCAACCGCACAGGATGTAAGCTGATCTGCCACCGGCTCA 118  
 Db 1350 GGAAGCGCTGTGATCTGATGATGCGGATGCTCGCGGCTTCAACGCGCGATCG 1409  
 QY 119 CCCCTACCTGGCGGAGGCGCAAGTGTCTGCTGACCTGGAACGAGAGAGTGTGACA 178  
 Db 1410 ATACCAACCTGGGCGCGGCGGACGACCGCTGCGCGGTGACATGAGCTGCGGTACA 1469  
 QY 179 GCGCCCGGCTCAACCAAGCGCATGCGGCGGGCACTTGTGATTTGTGAAGACTC 238  
 Db 1470 TCGCCACGTTGGCGCCCACTGAGGTGGCGAGCGCGATGCGGCGGTGAGCAAGCCT 1529  
 QY 239 GAGGGAAGA 248  
 Db 1530 GTTCGAAGA 1539

RESULT 12  
 US-09-252-991A-351/c  
 ; Sequence 351, Application US/09252991A  
 ; Patent No. 6551795

/ GENERAL INFORMATION:  
/ APPLICANT: Marc J. Rubenfield et al.  
/ TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS  
/ FILE REFERENCE: 107196.136  
/ CURRENT APPLICATION NUMBER: US/09/252,991A  
/ PRIOR FILING DATE: 1999-02-18  
/ PRIOR APPLICATION NUMBER: US 60/074,788  
/ PRIOR FILING DATE: 1998-02-18  
/ PRIOR APPLICATION NUMBER: US 60/094,190  
/ NUMBER OF SEQ ID NOS: 33142  
/ SEQ ID NO 351  
/ LENGTH: 2016  
/ TYPE: DNA  
/ ORGANISM: Pseudomonas aeruginosa  
US-09-252-991A-351

Query Match 10.3%; Score 42.8; DB 4; Length 2016;  
Best Local Similarity 51.6%; Pred. No. 0.18;  
Matches 98; Conservative 0; Mismatches 92; Indels 0; Gaps 0;

QY 59 GGGAGATGATGAGAGACCAACCGCACAGATGTAAGTCACTGCCCCAGCGGCTCA 118  
DB 623 GGAAGCGGCTGTGATGCTGATGCTCCGATTCCTCGCGGCTTCAACGGCCGATCG 564  
QY 119 CCCCCTACCTGCGCCAGCCAGAGTGTGCTGACGAGAGAGAGTGTGACA 178  
DB 563 ATACCAACGCGGCGCCGCGACGACGCGTGGCGGCTGAGATGAGCTGCGGTACA 504  
QY 179 GCCCCGGCTCACCAACAGCGCCATCGGCGGCGCACTTGTGATTGTAAGACTC 238  
DB 503 TCGCCACGTTGCGCCCGACGCTGAGGTGGCGAGGCGCGGCTGACCAACGCTC 444  
QY 239 GAGGAGAGAA 248  
DB 443 GTTCGAGAGA 434

## RESULT 13

US-09-252-991A-5873  
/ Sequence 5873, Application US/09252991A  
/ Patent No. 6551795  
/ GENERAL INFORMATION:  
/ APPLICANT: Marc J. Rubenfield et al.  
/ TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS  
/ FILE REFERENCE: 107196.136  
/ CURRENT APPLICATION NUMBER: US/09/252,991A  
/ PRIOR FILING DATE: 1999-02-18  
/ PRIOR APPLICATION NUMBER: US 60/074,788  
/ PRIOR FILING DATE: 1998-02-18  
/ PRIOR APPLICATION NUMBER: US 60/094,190  
/ NUMBER OF SEQ ID NOS: 33142  
/ SEQ ID NO 5873  
/ LENGTH: 534  
/ TYPE: DNA  
/ ORGANISM: Pseudomonas aeruginosa  
US-09-252-991A-5873

Query Match 10.2%; Score 42.6; DB 4; Length 534;  
Best Local Similarity 54.9%; Pred. No. 0.15;  
Matches 84; Conservative 0; Mismatches 69; Indels 0; Gaps 0;

QY 99 CATCTGCCCCAGCGGCTCAACCCCTACCTGCGCCAGCAAGTGTGTCAGCTGA 158  
DB 337 CAGCGCCGACCTGCTTCTCCGCGCCTTGCGCAAGCGGACGCTGCGCGCACGCC 396  
QY 159 CGAGAGAGAGTGTCTCACAGCCCGGCTCACCAACAGCGCCATGCGGCGCGGCACTT 218  
DB 397 GCTGAGAGAGGCGCTGAAAGGCGGTGCTTCAATGAGATGAGACGAGACGAGACTCTCT 456

QY 219 GCTGATTGTGTAAGACTCGAGGAGAACGG 251  
DB 457 GCTAGCGCGCGCGGCGCGCGCGCGCAAGGCGCG 489

## RESULT 14

US-09-252-991A-5829  
/ Sequence 5829, Application US/09252991A  
/ Patent No. 6551795  
/ GENERAL INFORMATION:  
/ APPLICANT: Marc J. Rubenfield et al.  
/ TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS  
/ FILE REFERENCE: 107196.136  
/ CURRENT APPLICATION NUMBER: US/09/252,991A  
/ PRIOR FILING DATE: 1999-02-18  
/ PRIOR APPLICATION NUMBER: US 60/074,788  
/ PRIOR FILING DATE: 1998-02-18  
/ PRIOR APPLICATION NUMBER: US 60/094,190  
/ NUMBER OF SEQ ID NOS: 33142  
/ SEQ ID NO 5829  
/ LENGTH: 1968  
/ TYPE: DNA  
/ ORGANISM: Pseudomonas aeruginosa  
US-09-252-991A-5829

Query Match 10.2%; Score 42.6; DB 4; Length 1968;  
Best Local Similarity 54.9%; Pred. No. 0.2;  
Matches 84; Conservative 0; Mismatches 69; Indels 0; Gaps 0;

QY 99 CATCTGCCCCAGCGGCTCAACCCCTACCTGCGCCAGGAGTGTGTCAGCTGA 158  
DB 1812 CAGCGCGACCTGCTTCTCCGCGCCTGCGCAAGCGGAGTGTGCGCGCACGCC 1871  
QY 159 CGAGAGAGAGTGTCTGACAGCCCGGCTGACCAACAGCGCCATGCGGCGCGGCACTT 218  
DB 1872 GCTGAGAGAGCGCTGAAAGCGGCTGCTCATGAGATGACGAGCCGAGCACTCTCT 1931  
QY 219 GCTGATTGTGTAAGACTCGAGGAGAACGG 251  
DB 1932 GCTAGCGCGCGCGGCGCGCGCGCGCAAGGCGCG 1964

## RESULT 15

US-09-252-991A-5779/c  
/ Sequence 5779, Application US/09252991A  
/ Patent No. 6551795  
/ GENERAL INFORMATION:  
/ APPLICANT: Marc J. Rubenfield et al.  
/ TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS  
/ FILE REFERENCE: 107196.136  
/ CURRENT APPLICATION NUMBER: US/09/252,991A  
/ PRIOR FILING DATE: 1999-02-18  
/ PRIOR APPLICATION NUMBER: US 60/074,788  
/ PRIOR FILING DATE: 1998-02-18  
/ PRIOR APPLICATION NUMBER: US 60/094,190  
/ NUMBER OF SEQ ID NOS: 33142  
/ SEQ ID NO 5779  
/ LENGTH: 2190  
/ TYPE: DNA  
/ ORGANISM: Pseudomonas aeruginosa  
US-09-252-991A-5779

Query Match 10.2%; Score 42.6; DB 4; Length 2190;  
Best Local Similarity 54.9%; Pred. No. 0.2;  
Matches 84; Conservative 0; Mismatches 69; Indels 0; Gaps 0;

QY 99 CATCTGCCCCAGCGGCTCAACCCCTACCTGCGCCAGGAGTGTGTCAGCTGA 158  
DB 160 CAGCGCGACCTGCTTCTCCGCGCCTGCGCAAGCGGACGCTGCTCCGCGCACGCC 101

QY 159 CGAGAGAGAGTGTGTGACAGACCCCGGCTCACCAACAGCGCCATGCGGCGCGGCACTT 218  
 Db 100 GCTGAGAGAGGCGCTTGAAAGCGGTCCCTTCATGAGATGACGACCGGACGACCTCCT 41  
 QY 219 GCTGATTGCTGAGACTCGAGGGAAGAACGG 251  
 Db 40 GCTGAGCGGCGCGGCGCGCGCGCAAGGCGG 8

Search completed: February 17, 2004, 23:58:57  
 Job time : 61.1645 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 17, 2004, 22:52:28 (Search time 269.576 Seconds  
(without alignments)  
5698.119 Million cell updates/sec

Title: US-10-032-159a-15

Sequence: 1 atgggggaactgtgcgcag.....ggctctcgttaaccacggn 417

Scoring table: IDENTITY\_NUC  
Gapop 10.0, Gapept 1.0

Searched: 244703 segs, 1841816367 residues

Total number of hits satisfying chosen parameters: 4899406

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database: Published Applications NA.\*

- 1: /cgn2\_6/ptodata/1/pubpna/US07\_PUBCOMB.seq.\*
- 2: /cgn2\_6/ptodata/1/pubpna/PCT\_NEW\_PUB.seq.\*
- 3: /cgn2\_6/ptodata/1/pubpna/US06\_NEW\_PUB.seq.\*
- 4: /cgn2\_6/ptodata/1/pubpna/US06\_PUBCOMB.seq.\*
- 5: /cgn2\_6/ptodata/1/pubpna/US07\_NEW\_PUB.seq.\*
- 6: /cgn2\_6/ptodata/1/pubpna/PCTUS\_PUBCOMB.seq.\*
- 7: /cgn2\_6/ptodata/1/pubpna/US08\_NEW\_PUB.seq.\*
- 8: /cgn2\_6/ptodata/1/pubpna/US08\_PUBCOMB.seq.\*
- 9: /cgn2\_6/ptodata/1/pubpna/US09\_PUBCOMB.seq.\*
- 10: /cgn2\_6/ptodata/1/pubpna/US09\_PUBCOMB.seq.\*
- 11: /cgn2\_6/ptodata/1/pubpna/US09C\_PUBCOMB.seq.\*
- 12: /cgn2\_6/ptodata/1/pubpna/US09\_NEW\_PUB.seq.\*
- 13: /cgn2\_6/ptodata/1/pubpna/US10\_NEW\_PUB.seq.\*
- 14: /cgn2\_6/ptodata/1/pubpna/US10\_PUBCOMB.seq.\*
- 15: /cgn2\_6/ptodata/1/pubpna/US10B\_PUBCOMB.seq.\*
- 16: /cgn2\_6/ptodata/1/pubpna/US10\_NEW\_PUB.seq.\*
- 17: /cgn2\_6/ptodata/1/pubpna/US60\_NEW\_PUB.seq.\*
- 18: /cgn2\_6/ptodata/1/pubpna/US60\_PUBCOMB.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	415	99.5	417	14	US-10-032-159a-15
2	415	99.5	3417	9	US-09-767-215-4
3	351	84.2	3012	9	US-09-767-215-3
4	351	84.2	3756	12	US-10-104-047-315
5	351	84.2	3951	9	US-09-767-215-1
6	276	66.2	276	14	US-10-032-159a-17
7	117.8	28.2	3096	11	US-09-798-412-9
8	117.8	28.2	3096	15	US-10-325-917-9
9	117.8	28.2	3949	11	US-09-798-412-7
10	117.8	28.2	3949	15	US-10-325-917-7
11	111.8	26.8	1141	14	US-10-032-159a-37
12	104	24.9	3441	11	US-09-798-412-12
13	104	24.9	3441	15	US-10-325-917-12
14	104	24.9	3744	14	US-10-032-159a-7
15	104	24.9	4276	11	US-09-798-412-10

16	104	24.9	4276	15	US-10-325-917-10	Sequence 10, Appl
17	100.4	24.1	276	14	US-10-032-159a-9	Sequence 9, Appl
18	94.4	22.6	1101	14	US-10-032-159a-1	Sequence 1, Appl
19	94.4	22.6	1608	11	US-09-798-412-6	Sequence 6, Appl
20	94.4	22.6	1608	15	US-10-325-917-6	Sequence 6, Appl
21	94.4	22.6	2098	11	US-09-798-412-4	Sequence 4, Appl
22	94.4	22.6	2098	15	US-10-325-917-4	Sequence 4, Appl
23	94.4	22.6	2176	14	US-10-032-159a-19	Sequence 19, Appl
24	94	22.5	514	10	US-09-798-692-3560	Sequence 3560, Ap
25	94	22.5	514	12	US-10-057-4758-3560	Sequence 3560, Ap
26	94	22.5	514	12	US-10-154-8848-3560	Sequence 3560, Ap
27	94	22.5	514	15	US-10-040-862-3560	Sequence 3560, Ap
28	85.2	20.4	1608	11	US-09-798-412-3	Sequence 3, Appl
29	85.2	20.4	1608	15	US-10-325-917-3	Sequence 3, Appl
30	85.2	20.4	1879	11	US-09-798-412-1	Sequence 1, Appl
31	85.2	20.4	1879	15	US-10-325-917-1	Sequence 1, Appl
32	69.8	16.7	216	14	US-10-032-159a-3	Sequence 3, Appl
33	51.8	12.4	281	14	US-10-032-159a-21	Sequence 21, Appl
34	51.8	12.4	500	13	US-10-029-386-12805	Sequence 12805, A
35	49.8	11.9	804	13	US-10-027-632-172688	Sequence 172688, A
36	49.8	11.9	804	14	US-10-027-632-172688	Sequence 172688, A
37	49	11.8	858	13	US-10-027-632-166977	Sequence 166977, A
38	49	11.8	858	14	US-10-029-386-26505	Sequence 26505, A
39	48	11.5	145	13	US-10-027-632-290760	Sequence 290760, A
40	47	11.3	509	13	US-10-027-632-290761	Sequence 290761, A
41	47	11.3	509	14	US-10-027-632-290760	Sequence 290760, A
42	47	11.3	509	14	US-10-027-632-290761	Sequence 290761, A
43	47	11.3	509	14	US-10-027-632-290761	Sequence 290761, A
44	45	10.8	36620	11	US-09-952-060-30	Sequence 30, Appl
45	44.6	10.7	1548	12	US-10-369-493-31289	Sequence 31289, A

ALIGNMENTS

RESULT 1  
US-10-032-159a-15  
Sequence 15, Application US/10032159A  
Publication No. US20020164703A1  
GENERAL INFORMATION:  
APPLICANT: Pawlowski, Krzysztof  
APPLICANT: Reed, John C.  
TITLE OF INVENTION: CARD-DOMAIN CONTAINING POLYPEPTIDES.  
TITLE OF INVENTION: ENCODING NUCLEIC ACIDS, AND METHODS OF USE  
FILE REFERENCE: P-LJ 5100  
CURRENT APPLICATION NUMBER: US/10/032,159A  
CURRENT FILING DATE: 2001-12-19  
PRIOR APPLICATION NUMBER: US 60/257,457  
PRIOR FILING DATE: 2000-12-21  
NUMBER OF SEQ ID NOS: 37  
SOFTWARE: FastSeq For Windows Version 4.0  
SEQ ID NO 15  
LENGTH: 417  
TYPE: DNA  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: CDS  
LOCATION: (1)...(417)  
FEATURE:  
NAME/KEY: misc feature  
LOCATION: 416..417  
OTHER INFORMATION: n = A,T,C or G  
US-10-032-159a-15  
Query Match 99.5% Score 415; DB 14; Length 417;  
Best Local Similarity 100.0%; Pred. No. 1.1e-107;  
Matches 415; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 ATGGGGAACTGTGCGGAGGACTCGCACTGACGAGCACTGACGAGAGACTGTGG 60  
DB 1 ATGGGGAACTGTGCGGAGGACTCGCACTGACGAGCACTGACGAGAGACTGTGG 60

```

QY 61 GAGATGATGAGAGAGCCACCGCCACAGATGTCAGCTGATCTGCCACGCGCTCACC 120
Db 61 GAGATGATGAGAGAGCCACCGCCACAGATGTCAGCTGATCTGCCACGCGCTCACC 120
QY 121 CCTACTGTCGCGCAGGCGCAAGGTGTGTGTCCAGCTGAGCAGAGAGAGAGTGTGTGACAGC 180
Db 121 CCTACTGTCGCGCAGGCGCAAGGTGTGTGTCCAGCTGAGCAGAGAGAGAGTGTGTGACAGC 180
QY 181 CCCCAGCTCACCAACAGCGCCATGCGGCGCGGCACTTGCTGATTTGCTGAAGAATCGA 240
Db 181 CCCCAGCTCACCAACAGCGCCATGCGGCGCGGCACTTGCTGATTTGCTGAAGAATCGA 240
QY 241 GGGAGAGAGCGGGCCATCGCTTCTCTGAGAGCTGGAAGTTCCAAACCTGACCTTAC 300
Db 241 GGGAGAGAGCGGGCCATCGCTTCTCTGAGAGCTGGAAGTTCCAAACCTGACCTTAC 300
QY 301 ACCCTGTCACCGGCGCTGAGGCTGATGTGTGACTTCAATTAAGGCTGAGAGCTCC 360
Db 301 ACCCTGTCACCGGCGCTGAGGCTGATGTGTGACTTCAATTAAGGCTGAGAGCTCC 360
QY 361 GACTTGACGCTTTGGCAGGCACTTCTAGGAACCTCAGGCTCTGTGTAACCCAG 415
Db 361 GACTTGACGCTTTGGCAGGCACTTCTAGGAACCTCAGGCTCTGTGTAACCCAG 415

```

## RESULT 2

```

US-09-767-215-4
; Sequence 4, Application US/09767215
; Patent No. US20020081636A1
; GENERAL INFORMATION:
; APPLICANT: Berlin, John
; TITLE OF INVENTION: NOVEL MOLECULES OF THE CARD-RELATED
; FILE REFERENCE: 07334-142001
; CURRENT APPLICATION NUMBER: US/09/767,215
; PRIOR FILING DATE: 2001-01-22
; PRIOR APPLICATION NUMBER: 60/181,159
; PRIOR FILING DATE: 2000-02-09
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 3417
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)...(3417)
US-09-767-215-4

```

```

Query Match 99.5%; Score 415; DB 9; Length 3417;
Best Local Similarity 100.0%; Pred. No. 1.5e-107; Indels 0; Gaps 0;
Matches 415; Conservative 0; Mismatches 0;

```

```

QY 1 ATGGGGGAATCTGTGCGCAGGAGACTCCGACTCAGGCGCACTGAGAGAGAGCACTGTGG 60
Db 1 ATGGGGGAATCTGTGCGCAGGAGACTCCGACTCAGGCGCACTGAGAGAGAGCACTGTGG 60
QY 61 GAGATGATGAGAGAGCCACCGCCACAGATGTAAGCTGATGTGCCCAAGCGGCTCAAC 120
Db 61 GAGATGATGAGAGAGCCACCGCCACAGATGTAAGCTGATGTGCCCAAGCGGCTCAAC 120
QY 121 CCTACTGTCGCGCAGGCGCAAGGTGTGTGTCCAGCTGAGCAGAGAGAGAGTGTGTGACAGC 180
Db 121 CCTACTGTCGCGCAGGCGCAAGGTGTGTGTCCAGCTGAGCAGAGAGAGAGTGTGTGACAGC 180
QY 181 CCCCAGCTCACCAACAGCGCCATGCGGCGCGGCACTTGCTGATTTGCTGAAGAATCGA 240
Db 181 CCCCAGCTCACCAACAGCGCCATGCGGCGCGGCACTTGCTGATTTGCTGAAGAATCGA 240
QY 241 GGGAGAGAGCGGGCCATCGCTTCTCTGAGAGCTGGAAGTTCCAAACCTGACCTTAC 300
Db 241 GGGAGAGAGCGGGCCATCGCTTCTCTGAGAGCTGGAAGTTCCAAACCTGACCTTAC 300

```

```

QY 301 ACCCTGTCACCGGCGCTGAGGCTGATGTGTGACTTCAATTAAGGCTGAGAGCTCC 360
Db 301 ACCCTGTCACCGGCGCTGAGGCTGATGTGTGACTTCAATTAAGGCTGAGAGCTCC 360
QY 361 GACTTGACGCTTTGGCAGGCACTTCTAGGAACCTCAGGCTCTGTGTAACCCAG 415
Db 361 GACTTGACGCTTTGGCAGGCACTTCTAGGAACCTCAGGCTCTGTGTAACCCAG 415

```

## RESULT 3

```

US-09-767-215-3
; Sequence 3, Application US/09767215
; Patent No. US20020081636A1
; GENERAL INFORMATION:
; APPLICANT: Berlin, John
; TITLE OF INVENTION: NOVEL MOLECULES OF THE CARD-RELATED
; FILE REFERENCE: 07334-142001
; CURRENT APPLICATION NUMBER: US/09/767,215
; PRIOR FILING DATE: 2001-01-22
; PRIOR APPLICATION NUMBER: 60/181,159
; PRIOR FILING DATE: 2000-02-09
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 3012
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-767-215-3

```

```

Query Match 84.2%; Score 351; DB 9; Length 3012;
Best Local Similarity 100.0%; Pred. No. 1.8e-89; Indels 0; Gaps 0;
Matches 351; Conservative 0; Mismatches 0;

```

```

QY 1 ATGGGGGAATCTGTGCGCAGGAGACTCCGACTCAGGCGCACTGAGAGAGAGCACTGTGG 60
Db 1 ATGGGGGAATCTGTGCGCAGGAGACTCCGACTCAGGCGCACTGAGAGAGAGCACTGTGG 60
QY 61 GAGATGATGAGAGAGCCACCGCCACAGATGTAAGCTGATGTGCCCAAGCGGCTCAAC 120
Db 61 GAGATGATGAGAGAGCCACCGCCACAGATGTAAGCTGATGTGCCCAAGCGGCTCAAC 120
QY 121 CCTACTGTCGCGCAGGCGCAAGGTGTGTGTCCAGCTGAGCAGAGAGAGTGTGTGACAGC 180
Db 121 CCTACTGTCGCGCAGGCGCAAGGTGTGTGTCCAGCTGAGCAGAGAGAGTGTGTGACAGC 180
QY 181 CCCCAGCTCACCAACAGCGCCATGCGGCGCGGCACTTGCTGATTTGCTGAAGAATCGA 240
Db 181 CCCCAGCTCACCAACAGCGCCATGCGGCGCGGCACTTGCTGATTTGCTGAAGAATCGA 240
QY 241 GGGAGAGAGCGGGCCATCGCTTCTCTGAGAGCTGGAAGTTCCAAACCTGACCTTAC 300
Db 241 GGGAGAGAGCGGGCCATCGCTTCTCTGAGAGCTGGAAGTTCCAAACCTGACCTTAC 300
QY 301 ACCCTGTCACCGGCGCTGAGGCTGATGTGTGACTTCAATTAAGGCTGAGAGCTCC 360
Db 301 ACCCTGTCACCGGCGCTGAGGCTGATGTGTGACTTCAATTAAGGCTGAGAGCTCC 360

```

## RESULT 4

```

US-10-104-047-315
; Sequence 315, Application US/10104047
; Publication No. US20030236392A1
; GENERAL INFORMATION:
; APPLICANT: HELIX RESEARCH INSTITUTE
; TITLE OF INVENTION: NO. US20030236392A1el full length cDNA
; FILE REFERENCE: HI-A0105
; CURRENT APPLICATION NUMBER: US/10/104,047
; PRIOR FILING DATE: 2002-03-25
; NUMBER OF SEQ ID NOS: 4096
; SOFTWARE: Patentm Ver. 2.1

```



```

; TITLE OF INVENTION: PROTEIN FAMILY AND USES THEREOF
; FILE REFERENCE: 07334-327001
; CURRENT APPLICATION NUMBER: US/09/798,412
; PRIOR FILING DATE: 2001-03-02
; PRIOR APPLICATION NUMBER: US 09/728,260
; PRIOR FILING DATE: 2000-12-01
; PRIOR APPLICATION NUMBER: US 09/685,791
; PRIOR FILING DATE: 2000-10-10
; PRIOR APPLICATION NUMBER: US 09/513,904
; PRIOR FILING DATE: 2000-02-25
; PRIOR APPLICATION NUMBER: US 09/507,533
; PRIOR FILING DATE: 2000-02-18
; PRIOR APPLICATION NUMBER: US 60/168,780
; PRIOR FILING DATE: 1999-12-03
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 9
; LENGTH: 3096
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-09-798-412-9

```

```

Query Match      28.2%; Score 117.8; DB 11; Length 3096;
Best Local Similarity 59.3%; Pred. No. 1.5e-23;
Matches 200; Conservative 0; Mismatches 137; Indels 0; Gaps 0;

```

```

QY 18 CAGGAGCTCCGCACTACAGGCACTGAGAGAGACACTGTGGAGATGAGAGCA 77
DB 42 CCGGGCCGGCTCGGGGTCTGAGGCGAGAGAGACCGCTGTGGAGGCAATCGAGGGGT 101
QY 78 CCGCACAAGATGTAAGCTGATGCTCCCAAGCCGCTCAACCCCTTACTGCGCCAGGC 137
DB 102 CCGGATCGGGCTGCTCCGCGCTGACCCGCAAGCTCAAGCCCTTACTGCGCCAGTG 161
QY 138 CAAGGTGCTGTGCGCACTGAGAGAGAGAGAGAGTGTGACAGCCCGGCTCAACCAAG 197
DB 162 CCGGGTCACTGACAGAGAGAGAGAGAGAGAGAGTGTGACAGCTTACCGCTTCCGTCGG 221
QY 198 CGCCATCGCGGCGCGGCACTTGTGATTTGCTGAGATCGAGAGAGAGAGAGAGGCGCAT 257
DB 222 CGTCAACCGCAGCCGCGCGCTGATGAGACATCTTGGCGCTGCGTGAAGAGGGGCTATGA 281
QY 258 CGCCTTCTGAGAGAGCCTGAAAGTTCCACAACTGACGTCTACACCTGTGACCGGGCT 317
DB 282 GGCCTTCTGGAAGCCCTGAGTTCTACTACCCGGAACACTTCAAGCTGCTCAAGGGCA 341
QY 318 GCAGCTGATGTGACTTCACTTCACTTAAGGGGTGAG 354
DB 342 GGAACCGCGCCAGCGCTGCTCATGATCCTCATGAG 378

```

```

RESULT 8
US-10-325-917-9
; Sequence 9, Application US/10325917
; Publication No. US20030113787A1
; GENERAL INFORMATION:
; APPLICANT: Bertin, John
; TITLE OF INVENTION: NOVEL MOLECULES OF THE CARD-RELATED
; FILE REFERENCE: 07334-327001
; CURRENT APPLICATION NUMBER: US/10/325,917
; PRIOR FILING DATE: 2002-12-20
; PRIOR APPLICATION NUMBER: US/09/798,412
; PRIOR FILING DATE: 2001-03-02
; PRIOR APPLICATION NUMBER: US 09/728,260
; PRIOR FILING DATE: 2000-12-01
; PRIOR APPLICATION NUMBER: US 09/685,791
; PRIOR FILING DATE: 2000-10-10
; PRIOR APPLICATION NUMBER: US 09/513,904
; PRIOR FILING DATE: 2000-02-25
; PRIOR APPLICATION NUMBER: US 09/507,533
; PRIOR FILING DATE: 2000-02-18
; PRIOR APPLICATION NUMBER: US 60/168,780

```

```

; PRIOR FILING DATE: 1999-12-03
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 9
; LENGTH: 3096
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-10-325-917-9

```

```

Query Match      28.2%; Score 117.8; DB 15; Length 3096;
Best Local Similarity 59.3%; Pred. No. 1.5e-23;
Matches 200; Conservative 0; Mismatches 137; Indels 0; Gaps 0;

```

```

QY 18 CAGGAGCTCCGCACTACAGGCACTGAGAGAGAGAGAGTGTGACAGCCCGGCTCAACCAAG 77
DB 42 CCGGGCCGGCTCGGGGTCTGAGGCGAGAGAGACCGCTGTGGAGGCAATCGAGGGGT 101
QY 78 CCGCACAAGATGTAAGCTGATGCTCCCAAGCCGCTCAACCCCTTACTGCGCCAGGC 137
DB 102 CCGGATCGGGCTGCTCCGCGCTGACCCGCAAGCTCAAGCCCTTACTGCGCCAGTG 161
QY 138 CAAGGTGCTGTGCGCACTGAGAGAGAGAGAGAGTGTGACAGCCCGGCTCAACCAAG 197
DB 222 CGTCAACCGCAGCCGCGCGCTGATGAGACATCTTGGCGCTGCGTGAAGAGGGGCTATGA 281
QY 258 CGCCTTCTGAGAGAGCCTGAAAGTTCCACAACTGACGTCTACACCTGTGACCGGGCT 317
DB 282 GGCCTTCTGGAAGCCCTGAGTTCTACTACCCGGAACACTTCAAGCTGCTCAAGGGCA 341
QY 318 GCAGCTGATGTGACTTCACTTCACTTAAGGGGTGAG 354
DB 342 GGAACCGCGCCAGCGCTGCTCATGATCCTCATGAG 378

```

```

RESULT 9
US-09-798-412-7
; Sequence 7, Application US/09798412
; Publication No. US20030109428A1
; GENERAL INFORMATION:
; APPLICANT: Bertin, John
; TITLE OF INVENTION: NOVEL MOLECULES OF THE CARD-RELATED
; FILE REFERENCE: 07334-327001
; CURRENT APPLICATION NUMBER: US/09/798,412
; PRIOR FILING DATE: 2001-03-02
; PRIOR APPLICATION NUMBER: US 09/728,260
; PRIOR FILING DATE: 2000-12-01
; PRIOR APPLICATION NUMBER: US 09/685,791
; PRIOR FILING DATE: 2000-10-10
; PRIOR APPLICATION NUMBER: US 09/513,904
; PRIOR FILING DATE: 2000-02-25
; PRIOR APPLICATION NUMBER: US 09/507,533
; PRIOR FILING DATE: 1999-12-03
; PRIOR APPLICATION NUMBER: US 60/168,780

```

```

; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 7
; LENGTH: 3949
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (41)...(3136)
; US-09-798-412-7

```

```

Query Match      28.2%; Score 117.8; DB 11; Length 3949;
Best Local Similarity 59.3%; Pred. No. 1.6e-23;
Matches 200; Conservative 0; Mismatches 137; Indels 0; Gaps 0;

```

QY 18 CAGGACTCCGCACTGACGCGCACTGAGACGAGACACTGTGGAGATGAGAGCA 77  
 DB 82 CCGGGCCGGCTCGGGCTCTGAGCGAGAGACGCGCTGTGGAGCAATGAGGGCGT 141  
 QY 78 CCGCCACAGAGATGTAAGCTGATCTGCCCCAGCCGCTCACCCCTTACCTGGCCAGAGC 137  
 DB 142 CCGGCAATCGGCTGGCTGCGCCCTGAAACCGGCGCAAGCTCAAGCCCTATCTGGCCAGT 201  
 QY 138 CAGGTGCTGTGCACTGAGACGAGAGAGAGTGTGCAAGCCCCGGCTCAACAAAG 197  
 DB 202 CCGGCTATGACGAGAGAGAGAGAGAGTGTGAGACTTACCGCTTCCCGCGCG 261  
 QY 198 CCGCATGCGGGCCGGGCACTTGTCTGTAAGTCTGAGAGTCCGAGGAAAGACGGGGCAT 257  
 DB 262 CGTCAACCGACCGGGGCGCTGATGACATCTTGGCGCTGGCGCAAGAGGGGCTATGA 321  
 QY 258 CCGCTTCTGAGAGAGCGCTGATTCACCAACCGCTGACGCTGATCACCGGGCT 317  
 DB 322 GCGCTTCTGAGAGCGCTTGAAGTCTTACTACCCGAAACATTCACGCTGCTACCGGGCA 381  
 QY 318 GCAGCTGATGTTGACTTCACTTAACTTAAAGCGGTGAG 354  
 DB 382 GGAACCGCGCCAGCGCTGCTCCATGATCTTCATGAG 418

RESULT 10  
 US-10-325-917-7  
 ; Sequence 7, Application US/10325917  
 ; Publication No. US20030113787A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Bertin, John  
 ; TITLE OF INVENTION: NOVEL MOLECULES OF THE CARD-RELATED  
 ; FILE OF INVENTION: PROTEIN FAMILY AND USES THEREOF  
 ; FILE REFERENCE: 07334-327001  
 ; CURRENT APPLICATION NUMBER: US/10/325,917  
 ; CURRENT FILING DATE: 2002-12-20  
 ; PRIOR APPLICATION NUMBER: US/09/798,412  
 ; PRIOR FILING DATE: 2001-03-02  
 ; PRIOR APPLICATION NUMBER: US 09/728,260  
 ; PRIOR FILING DATE: 2000-12-01  
 ; PRIOR APPLICATION NUMBER: US 09/685,791  
 ; PRIOR FILING DATE: 2000-10-10  
 ; PRIOR APPLICATION NUMBER: US 09/513,904  
 ; PRIOR FILING DATE: 2000-02-25  
 ; PRIOR APPLICATION NUMBER: US 09/507,533  
 ; PRIOR FILING DATE: 2000-02-18  
 ; PRIOR APPLICATION NUMBER: US 60/168,780  
 ; PRIOR FILING DATE: 1999-12-03  
 ; NUMBER OF SEQ ID NOS: 19  
 ; SOFTWARE: FastSeq for Windows Version 4.0  
 ; SEQ ID NO 7  
 ; LENGTH: 3949  
 ; TYPE: DNA  
 ; ORGANISM: Homo sapiens  
 ; FEATURES:  
 ; NAME/KEY: CDS  
 ; LOCATION: (41)...(3136)  
 ; US-10-325-917-7

Query Match 28.2%; Score 117.8; DB 15; Length 3949;  
 Best Local Similarity 59.3%; Pred. No. 1.6e-23;  
 Matches 200; Conservative 0; Mismatches 137; Indels 0; Gaps 0;  
 QY 18 CAGGACTCCGCACTGACGCGCACTGAGACGAGACACTGTGGAGATGAGAGCA 77  
 DB 82 CCGGGCCGGCTCGGGCTCTGAGCGAGAGACGCGCTGTGGAGCAATGAGGGCGT 141  
 QY 78 CCGCCACAGAGATGTAAGCTGATCTGCCCCAGCCGCTCACCCCTTACCTGGCCAGAGC 137  
 DB 142 CCGGCAATCGGCTGGCTGCGCCCTGAAACCGGCGCAAGCTCAAGCCCTATCTGGCCAGT 201  
 QY 138 CAGGTGCTGTGCACTGAGACGAGAGAGAGTGTGCAAGCCCCGGCTCAACAAAG 197

DB 202 CCGGCTATGACAGAGACGAGAGAGAGTGTGAGACCTTACCGCTTCCGCTGCG 261  
 QY 198 CCGCATGCGGGCCCGGCACTTGTGATTTGTGTAAGCTGAGAGAAAGAGGGCCAT 257  
 DB 262 CGTCAACCGCACCGGGCGCTGATGAGACATCTTGCCTGCGTGGCAAGAGGGCTATGA 321  
 QY 258 CCGCTTCTGAGAGAGCTTGAAGTTCACAAACCTGAGAGCTTACACCTTGTGACCGGGCT 317  
 DB 322 GCGCTTCTGAGAGCGCTTGAAGTTCCTACTACCCGAAACATTCACGCTGCTACGGGGCA 381  
 QY 318 GCAGCTGATGTTGACTTCACTTAACTTAAAGCGGTGAG 354  
 DB 382 GGAACCGCGCCAGCGCTGCTCCATGATCTTCATGAG 418

RESULT 11  
 US-10-032-159A-37  
 ; Sequence 37, Application US/10032159A  
 ; Publication No. US20020164703A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Pawlowski, Krzysztof  
 ; APPLICANT: Reed, John C.  
 ; APPLICANT: Godzik, Adam  
 ; TITLE OF INVENTION: CARD-DOMAIN CONTAINING POLYPEPTIDES,  
 ; FILE OF INVENTION: ENCODING NUCLEIC ACIDS, AND METHODS OF USE  
 ; FILE REFERENCE: P-LJ 5100  
 ; CURRENT APPLICATION NUMBER: US/10/032,159A  
 ; CURRENT FILING DATE: 2001-12-19  
 ; PRIOR APPLICATION NUMBER: US 60/257,457  
 ; PRIOR FILING DATE: 2000-12-21  
 ; NUMBER OF SEQ ID NOS: 37  
 ; SOFTWARE: FastSeq for Windows Version 4.0  
 ; SEQ ID NO 37  
 ; LENGTH: 1141  
 ; TYPE: DNA  
 ; ORGANISM: Homo sapiens  
 ; US-10-032-159A-37

Query Match 26.8%; Score 111.8; DB 14; Length 1141;  
 Best Local Similarity 94.3%; Pred. No. 6.7e-22;  
 Matches 116; Conservative 0; Mismatches 7; Indels 0; Gaps 0;  
 QY 1 ATGGGGAACTGTGCGGAGGAGCTCCGCACTGACGCGCACTGAGAGAGCACTGTGG 60  
 DB 161 ATGGGGAACTGTGCGGAGGAGCTCCGCACTGACGCGCACTGAGAGAGCACTGTGG 220  
 QY 61 GAGATGATGAGAGCCACCGGCAAGATCGTACGCTGATCTGCCCGGCGCTCAC 120  
 DB 221 GAGATGATGAGAGCCACCGGCAAGATCGTACGCTGATCTGCCCGGCAAGATTA 280  
 QY 121 CCC 123  
 DB 281 CCC 283

RESULT 12  
 US-09-798-412-12  
 ; Sequence 12, Application US/09798412  
 ; Publication No. US20030109428A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Bertin, John  
 ; TITLE OF INVENTION: NOVEL MOLECULES OF THE CARD-RELATED  
 ; FILE OF INVENTION: PROTEIN FAMILY AND USES THEREOF  
 ; FILE REFERENCE: 07334-327001  
 ; CURRENT APPLICATION NUMBER: US/09/798,412  
 ; CURRENT FILING DATE: 2001-03-02  
 ; PRIOR APPLICATION NUMBER: US 09/728,260  
 ; PRIOR FILING DATE: 2000-12-01  
 ; PRIOR APPLICATION NUMBER: US 09/685,791  
 ; PRIOR FILING DATE: 2000-10-10  
 ; PRIOR APPLICATION NUMBER: US 09/513,904  
 ; PRIOR FILING DATE: 2000-02-25

PRIOR APPLICATION NUMBER: US 09/507,533  
 PRIOR FILING DATE: 2000-02-18  
 PRIOR APPLICATION NUMBER: US 60/168,780  
 PRIOR FILING DATE: 1999-12-03  
 NUMBER OF SEQ ID NOS: 19  
 SOFTWARE: FastSeq for Windows Version 4.0  
 SEQ ID NO: 12  
 LENGTH: 3441  
 TYPE: DNA  
 ORGANISM: Homo sapiens  
 US-09-798-412-12

Query Match 24.9%; Score 104; DB 11; Length 3441;  
 Best Local Similarity 58.3%; Pred. No. 1.2e-19;  
 Matches 182; Conservative 0; Mismatches 130; Indels 0; Gaps 0;

QY 43 GACGAGAGACACTGTGGAGATGATGAGAGCCACCGGACAGATGTAAGCTGATC 102  
 DB 31 GAAGAGAGACGCTTGTGGAGATGATGAGATGTAAGCTGATC 90  
 QY 103 TGCCCCAGCGGCTTACCCCTTACCTGCGCCAGGCAAGTGTGCTGCGAGCTGACGAG 162  
 DB 91 AACCTGCCAAGCTACGCGCTTACCTGCGCTGAGTGAAGTATGATGAGAGATGAA 150  
 QY 163 GAGAGAGTGTGACAGCCCGGCTCACCAACAGCGCCATGCGGCGGCACTTGCTG 222  
 DB 151 GATGAAGTGTATATGCTTATGCTGCTCATCAAGATCAACGAGAGCGGCTGTG 210  
 QY 223 GATTGCTGAAGCTGAGGAGAAAGGAGGCGCATGCTTCTTGAGAGCTGAAGTTC 282  
 DB 211 GACATTTCACTACCAAGGGGCAAGGGGCTATGATGTCTTCTTGAGAGCTTAAGTTT 270  
 QY 283 CACAACCTGACCTTACACCTGCTCACCGGCTGCAAGCTGATTTACTTCACTGATAC 342  
 DB 271 TATTACCAAGACTGTACAACTGTGACTGGGAAAGAGCCCACTCGAGATTCTCCACC 330  
 QY 343 TTAGCGGTGAG 354  
 DB 331 ATTGTGTGTGAG 342

RESULT 13  
 US-10-325-917-12  
 Sequence 12, Application US/10325917  
 Publication No. US20030113787A1  
 GENERAL INFORMATION:  
 APPLICANT: Bertlin, John  
 TITLE OF INVENTION: NOVEL MOLECULES OF THE CARD-RELATED  
 TITLE OF INVENTION: PROTEIN FAMILY AND USES THEREOF  
 FILE REFERENCE: 07334-327001  
 CURRENT APPLICATION NUMBER: US/10/325,917  
 CURRENT FILING DATE: 2002-12-20  
 PRIOR APPLICATION NUMBER: US/09/798,412  
 PRIOR FILING DATE: 2001-03-02  
 PRIOR APPLICATION NUMBER: US 09/728,260  
 PRIOR FILING DATE: 2000-12-01  
 PRIOR APPLICATION NUMBER: US 09/685,791  
 PRIOR FILING DATE: 2000-10-10  
 PRIOR APPLICATION NUMBER: US 09/513,904  
 PRIOR FILING DATE: 2000-02-25  
 PRIOR APPLICATION NUMBER: US 09/507,533  
 PRIOR FILING DATE: 2000-02-18  
 PRIOR APPLICATION NUMBER: US 60/168,780  
 PRIOR FILING DATE: 1999-12-03  
 NUMBER OF SEQ ID NOS: 19  
 SOFTWARE: FastSeq for Windows Version 4.0  
 SEQ ID NO: 12  
 LENGTH: 3441  
 TYPE: DNA  
 ORGANISM: Homo sapiens  
 US-10-325-917-12

Query Match 24.9%; Score 104; DB 15; Length 3441;

Best Local Similarity 58.3%; Pred. No. 1.2e-19;  
 Matches 182; Conservative 0; Mismatches 130; Indels 0; Gaps 0;

QY 43 GACGAGAGACACTGTGGAGATGATGAGAGCCACCGGACAGATGTAAGCTGATC 102  
 DB 31 GAAGAGAGACGCTTGTGGAGATGATGAGATGTAAGCTGATC 90  
 QY 103 TGCCCCAGCGGCTTACCCCTTACCTGCGCCAGGCAAGTGTGCTGCGAGCTGACGAG 162  
 DB 91 AACCTGCCAAGCTACGCGCTTACCTGCGCTGAGTGAAGTATGATGAGAGATGAA 150  
 QY 163 GAGAGAGTGTGACAGCCCGGCTCACCAACAGCGCCATGCGGCGGCACTTGCTG 222  
 DB 151 GATGAAGTGTATATGCTTATGCTGCTCATCAAGATCAACGAGAGCGGCTGTG 210  
 QY 223 GATTGCTGAAGCTGAGGAGAAAGGAGGCGCATGCTTCTTGAGAGCTGAAGTTC 282  
 DB 211 GACATTTCACTACCAAGGGGCAAGGGGCTATGATGTCTTCTTGAGAGCTTAAGTTT 270  
 QY 283 CACAACCTGACCTTACACCTGCTCACCGGCTGCAAGCTGATTTACTTCACTGATAC 342  
 DB 271 TATTACCAAGACTGTACAACTGTGACTGGGAAAGAGCCCACTCGAGATTCTCCACC 330  
 QY 343 TTAGCGGTGAG 354  
 DB 331 ATTGTGTGTGAG 342

RESULT 14  
 US-10-032-159A-7  
 Sequence 7, Application US/10032159A  
 Publication No. US20020164703A1  
 GENERAL INFORMATION:  
 APPLICANT: Pawlowski, Krzysztof  
 APPLICANT: Reed, John C.  
 TITLE OF INVENTION: CARD-DOMAIN CONTAINING POLYPEPTIDES,  
 TITLE OF INVENTION: ENCODING NUCLEIC ACIDS, AND METHODS OF USE  
 FILE REFERENCE: P-1J 5100  
 CURRENT APPLICATION NUMBER: US/10/032,159A  
 CURRENT FILING DATE: 2001-12-19  
 PRIOR APPLICATION NUMBER: US 60/257,457  
 PRIOR FILING DATE: 2000-12-21  
 NUMBER OF SEQ ID NOS: 37  
 SOFTWARE: FastSeq for Windows Version 4.0  
 SEQ ID NO: 7  
 LENGTH: 3744  
 TYPE: DNA  
 ORGANISM: Homo sapiens  
 NAME/KEY: CDS  
 LOCATION: (1)...(3744)  
 US-10-032-159A-7

Query Match 24.9%; Score 104; DB 14; Length 3744;  
 Best Local Similarity 58.3%; Pred. No. 1.2e-19;  
 Matches 182; Conservative 0; Mismatches 130; Indels 0; Gaps 0;

QY 43 GACGAGAGACACTGTGGAGATGATGAGAGCCACCGGACAGATGTAAGCTGATC 102  
 DB 31 GAAGAGAGACGCTTGTGGAGATGATGAGATGTAAGCTGATC 90  
 QY 103 TGCCCCAGCGGCTTACCCCTTACCTGCGCCAGGCAAGTGTGCTGCGAGCTGACGAG 162  
 DB 91 AACCTGCCAAGCTACGCGCTTACCTGCGCTGAGTGAAGTATGATGAGAGATGAA 150  
 QY 163 GAGAGAGTGTGACAGCCCGGCTCACCAACAGCGCCATGCGGCGGCACTTGCTG 222  
 DB 151 GATGAAGTGTATATGCTTATGCTGCTCATCAAGATCAACGAGAGCGGCTGTG 210  
 QY 223 GATTGCTGAAGCTGAGGAGAAAGGAGGCGCATGCTTCTTGAGAGCTGAAGTTC 282  
 DB 211 GACATTTCACTACCAAGGGGCAAGGGGCTATGATGTCTTCTTGAGAGCTTAAGTTT 270



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 18, 2004, 03:27:47 ; Search time 74 Seconds  
(without alignments)  
298.149 Million cell updates/sec

Title: US-10-032-159a-16  
Perfect score: 720

Sequence: 1 MSGLCRDRLALTRLDSEITLW.....SPFDGLAGSRNRLILVTPX 139

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1107863 seqs, 158726573 residues

Total number of hits satisfying chosen parameters: 1107863

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%  
Listing first 45 summaries

Database :

1: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1980.DAT:\*  
2: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1981.DAT:\*  
3: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1982.DAT:\*  
4: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1983.DAT:\*  
5: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1984.DAT:\*  
6: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1985.DAT:\*  
7: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1986.DAT:\*  
8: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1987.DAT:\*  
9: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1988.DAT:\*  
10: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1989.DAT:\*  
11: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1990.DAT:\*  
12: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1991.DAT:\*  
13: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1992.DAT:\*  
14: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1993.DAT:\*  
15: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1994.DAT:\*  
16: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1995.DAT:\*  
17: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1996.DAT:\*  
18: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1997.DAT:\*  
19: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1998.DAT:\*  
20: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1999.DAT:\*  
21: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA2000.DAT:\*  
22: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA2001.DAT:\*  
23: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA2002.DAT:\*  
24: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA2003.DAT:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	719	99.9	139	24	ABG76062 Human caspase recr
2	719	99.9	139	22	AAE07165 Human predicted ca
3	617.5	85.6	1004	22	AAE07164 Human caspase recr
4	270.5	37.6	1147	22	AAU01207 Human caspase recr
5	270.5	37.6	1147	23	AAU01207 Human CARD-11. Ho
6	270.5	37.6	1247	24	ABG76061 Human caspase recr
7	246	34.2	1032	22	AAU01206 Human caspase recr
8	246	34.2	1032	23	AAU01206 Human CARD-10. Ho
9	246	34.2	1032	23	AAU73247 Human plakoglobin

10	223.5	31.0	536	22	AAU01204
11	223.5	31.0	536	22	AAU79552
12	215.5	29.9	366	22	AAE95617
13	215.5	29.9	366	24	ABG76060
14	215.5	29.9	536	22	AAU01205
15	215.5	29.9	536	23	AAU79553
16	209	29.0	174	21	AAE41067
17	209	29.0	174	23	ABP10782
18	131	18.2	1139	22	AAU30127
19	100	13.9	237	23	AAU73245
20	83	11.5	519	21	AAE59412
21	80	11.1	519	21	AAE44570
22	79	11.0	519	21	AAE44565
23	79	11.0	519	21	AAE44569
24	79	11.0	519	21	AAE44571
25	79	11.0	519	21	AAE44572
26	79	11.0	519	24	ABP58351
27	78	10.8	233	21	AAE59413
28	74.5	10.3	269	22	ABG02857
29	72.5	10.1	706	23	ABP40063
30	72.5	10.1	845	22	ABE62651
31	72	10.0	299	23	AAO14117
32	71.5	9.9	260	22	ABE50208
33	71.5	9.9	660	23	ABE32605
34	71.5	9.9	707	22	AAE33055
35	71.5	9.9	889	23	AAE20268
36	71.5	9.9	956	22	AAE33299
37	71.5	9.9	956	23	ABP62970
38	70.5	9.8	1266	23	AAE22544
39	69.5	9.7	352	21	AAE07820
40	69.5	9.7	497	23	ABE88889
41	69	9.6	203	20	AAE34778
42	69	9.6	361	21	AAE11457
43	69	9.6	361	21	AAE39266
44	69	9.6	371	21	AAE11456
45	69	9.6	371	21	AAE39265

## ALIGNMENTS

RESULT 1	ABG76062	standard; protein, 139 AA.
ID	ABG76062	
AC	ABG76062	
XX		
DT	09-MAY-2003	(first entry)
DE	Human caspase recruitment domain containing protein, CARD-12X.	
XX		
KW	CARD; caspase recruitment domain; apoptosis; cell adhesion; inflammation; cytokine receptor signalling; cancer; glioma; carcinoma; adenocarcinoma; CYC-containing polypeptide associated disorder; sarcoma; melanoma; hamartoma; leukaemia; lymphoma; keratinocyte hyperplasia; neoplasia; keloid; benign prostatic hyperplasia; fibrosis; restenosis; allergy; arthritis; lupus; Sjogren's syndrome; sepsis; human; Crohn's disease; ulcerative colitis; graft versus host disease; stroke; abnormal cell death disease; myocardial infarction; heart failure; neurodegenerative disease; Parkinson's disease; Alzheimer's disease; HIV; CARD-12X; caspase activator; caspase inhibitor.	
XX		
OS	Homo sapiens.	
XX		
Key	Location/Qualifiers	
FT	16..107	
FT	Domain	
FT	/label= CARD	
FT	/note= "Caspase recruitment domain. Specifically claimed in claim 12"	
FT	Misc-difference 139	
FT	/label= Unknown	
FT	/note= "Encoded by GNN"	
XX		



PN US2002164703-A1.  
 XX 07-NOV-2002.  
 XX 19-DEC-2001; 2001US-0032159.  
 XX 21-DEC-2000; 2000US-257457P.  
 XX  
 PA (PAWL/) PAWLOWSKI K.  
 PA (REED/) REED J C.  
 PA (GODZ/) GODZIK A.  
 XX  
 PI Pawlowaki K, Reed JC, Godzik A;  
 XX  
 DR WPI; 2003-268137/28.  
 DR N-PSDB; ABX11431.  
 XX  
 PT New isolated CARD-containing nucleic acids, useful for the diagnosis  
 PT and treatment of disorders with aberrant expression or activity of the  
 PT CARD-containing polypeptide, such as cancer, stroke, arthritis, heart  
 PT failure and AIDS  
 XX  
 PS Claim 11; Fig 3; 34pp; English.  
 XX  
 CC The invention relates to an isolated nucleic acid molecule encoding a  
 CC caspase recruitment domain (CARD) containing polypeptide. CARD containing  
 CC polypeptides are involved in apoptosis (as caspase activators and caspase  
 CC inhibitors), cell adhesion, inflammation and cytokine receptor  
 CC signalling. The methods and compositions of the present invention are  
 CC useful for the diagnosis and treatment of disorders associated with the  
 CC aberrant expression or activity of the CARD containing polypeptide such  
 CC as cancer, glioma, carcinoma, adenocarcinoma, sarcoma, melanoma, keloid,  
 CC hamartoma, leukaemia, lymphoma, keratinocyte hyperplasia, neoplasia,  
 CC benign prostatic hypertrophy, inflammatory hyperplasia, fibrosis,  
 CC restenosis, allergies, inflammatory diseases such as arthritis, lupus,  
 CC Sjogren's syndrome, Crohn's disease, ulcerative colitis, graft versus  
 CC host disease, sepsis, abnormal cell death diseases such as stroke,  
 CC myocardial infarction, heart failure, neurodegenerative diseases like  
 CC Parkinson's disease and Alzheimer's disease, and HIV infection. The  
 CC present sequence represents the amino acid sequence of the human  
 CC caspase recruitment domain containing protein, CARD-12X.  
 CC  
 SQ Sequence 139 AA;  
 Query Match 99.9%; Score 719; DB 24; Length 139;  
 Best Local Similarity 100.0%; Pred. No. 4.1e-82;  
 Matches 138; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 MGEICRDSALTLADEETLWEMESHRRIVRCICPSRLTPYLROAKYLQDDEEVLTHS 60  
 DB 1 MGEICRDSALTLADEETLWEMESHRRIVRCICPSRLTPYLROAKYLQDDEEVLTHS 60  
 QY 61 PRLTNSAMRAGHLIDLKTRGKNGAIAFLESLEFHPNDVYTLVTGLOPDVDFSNFGSSS 120  
 DB 61 PRLTNSAMRAGHLIDLKTRGKNGAIAFLESLEFHPNDVYTLVTGLOPDVDFSNFGSSS 120  
 QY 121 DFDGLAGTSRNLRLVTP 138  
 DB 121 DFDGLAGTSRNLRLVTP 138  
 DB 121 DFDGLAGTSRNLRLVTP 138  
 RESULT 2  
 ID AAE07165 standard; Protein; 1139 AA.  
 XX AAE07165;  
 AC  
 XX 06-NOV-2001 (first entry)  
 XX Human predicted caspase recruitment domain-14 (CARD-14).  
 XX Human; caspase recruitment domain-14; CARD-14; chromosome 17;  
 KM nuclear factor-kappa B; NF-kB; cell growth; cell death; cancer; therapy;

KM autoimmune disorder; systemic lupus erythematosus; neurological disorder;  
 KM Alzheimer's disease; Parkinson's disease; inflammatory disorder; anaemia;  
 KM haematological disorder; myelodysplastic syndrome; myocardial infarction;  
 KM stroke; immune disorder; Crohn's disease; allergic rhinitis; infection;  
 KM cell signalling disorder; cytostatic; immunosuppressive; nootropic;  
 KM neuroprotective; antiviral; antibacterial.  
 XX  
 OS Homo sapiens.  
 XX  
 FH Key Location/Qualifiers  
 FT Misc-difference 700  
 FT /note= "Encoded by TGS"  
 XX  
 XX WO200159065-A2.  
 XX 16-AUG-2001.  
 XX  
 XX 22-JAN-2001; 2001WO-US02087.  
 XX  
 XX 09-FEB-2000; 2000US-0181159.  
 XX  
 XX (MILL-) MILLENNIUM PHARM INC.  
 XX  
 PI Bertin J;  
 XX  
 DR WPI; 2001-497073/54.  
 DR N-PSDB; AAD13448.  
 XX  
 PT An isolated caspase recruitment domain polypeptide useful for  
 PT regulating growth and cell death and useful for the treatment of cancer  
 PT  
 PS Disclosure; Fig 2A-2C; 109pp; English.  
 XX  
 CC The present sequence is predicted human caspase recruitment domain-14  
 CC (CARD-14). The CARD-14 gene is located on chromosome 17. The CARD-14 is  
 CC used for the detection of modulators that modulates the ability of  
 CC CARD-14 to bind to Bcl-10 and stimulates phosphorylation of Bcl-10 or  
 CC activation of nuclear factor-kappa B (NF-kB). The CARD-14 is useful for  
 CC regulating growth and cell death and useful for the treatment of cancer.  
 CC It is also useful for the treatment of autoimmune disorders (e.g.,  
 CC systemic lupus erythematosus), neurological disorders (e.g., Alzheimer's  
 CC and Parkinson's disease, inflammatory disorders, haematological disorders  
 CC (e.g., anaemia, myelodysplastic syndromes), myocardial infarctions,  
 CC strokes, immune disorders (e.g., Crohn's disease, allergic rhinitis),  
 CC cell signalling disorders and certain viral and bacterial infections.  
 CC  
 SQ Sequence 1139 AA;  
 Query Match 99.9%; Score 719; DB 22; Length 1139;  
 Best Local Similarity 100.0%; Pred. No. 9.5e-81;  
 Matches 138; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 MGEICRDSALTLADEETLWEMESHRRIVRCICPSRLTPYLROAKYLQDDEEVLTHS 60  
 DB 1 MGEICRDSALTLADEETLWEMESHRRIVRCICPSRLTPYLROAKYLQDDEEVLTHS 60  
 QY 61 PRLTNSAMRAGHLIDLKTRGKNGAIAFLESLEFHPNDVYTLVTGLOPDVDFSNFGSSS 120  
 DB 61 PRLTNSAMRAGHLIDLKTRGKNGAIAFLESLEFHPNDVYTLVTGLOPDVDFSNFGSSS 120  
 QY 121 DFDGLAGTSRNLRLVTP 138  
 DB 121 DFDGLAGTSRNLRLVTP 138  
 DB 121 DFDGLAGTSRNLRLVTP 138  
 RESULT 3  
 ID AAE07164 standard; Protein; 1004 AA.  
 XX AAE07164;  
 AC  
 XX 06-NOV-2001 (first entry)

XX Human caspase recruitment domain-14 (CARD-14).  
XX Human, caspase recruitment domain-14; CARD-14; chromosome 17;  
XX nuclear factor-kappa B; NF-kB; cell growth; cell death; cancer; therapy;  
KW autoimmune disorder; systemic lupus erythematosus; neurological disorder;  
KW Alzheimer's disease; Parkinson's disease; inflammatory disorder; anaemia;  
KW haematological disorder; myelodysplastic syndrome; myocardial infarction;  
KW stroke; immune disorder; Crohn's disease; allergic rhinitis; infection;  
KW cell signalling disorder; cytostatic; immunosuppressive; noctropic;  
KW neuroprotective; antiviral; antibacterial.  
XX Homo sapiens.  
XX  
XX  
XX Key Location/Qualifiers  
FT Modified-site /note= "cAMP- and cGMP-dependent protein kinase  
phosphorylation site"  
FT Domain 10..116  
FT /label= CARD\_domain  
FT Modified-site 12..15  
FT /note= "Casein kinase II phosphorylation site"  
FT Modified-site 18..21  
FT /note= "Casein kinase II phosphorylation site"  
FT Modified-site 25..27  
FT /note= "Protein kinase C phosphorylation site"  
FT Modified-site 60..62  
FT /note= "Protein kinase C phosphorylation site"  
FT Modified-site 91..93  
FT /note= "Protein kinase C phosphorylation site"  
FT Modified-site 114..117  
FT /note= "N-glycosylation site"  
FT Modified-site 117..122  
FT /note= "N-myristoylation site"  
FT Modified-site 121..123  
FT /note= "Protein kinase C phosphorylation site"  
FT Domain 126..420  
FT /label= Coiled\_Coil\_domain  
FT Modified-site 130..135  
FT /note= "N-myristoylation site"  
FT Modified-site 134..137  
FT /note= "Casein kinase II phosphorylation site"  
FT Modified-site 161..166  
FT /note= "N-myristoylation site"  
FT Modified-site 165..168  
FT /note= "Casein kinase II phosphorylation site"  
FT Modified-site 220..227  
FT /note= "Tyrosine kinase phosphorylation site"  
FT Modified-site 221..224  
FT /note= "Casein kinase II phosphorylation site"  
FT Domain 239..325  
FT /label= K-Box\_domain  
FT Modified-site 240..243  
FT /note= "Casein kinase II phosphorylation site"  
FT Modified-site 250..252  
FT /note= "Protein kinase C phosphorylation site"  
FT Modified-site 253..256  
FT /note= "Casein kinase II phosphorylation site"  
FT Modified-site 259..262  
FT /note= "Casein kinase II phosphorylation site"  
FT Modified-site 280..283  
FT /note= "Casein kinase II phosphorylation site"  
FT Modified-site 290..293  
FT /note= "Casein kinase II phosphorylation site"  
FT Modified-site 297..300  
FT /note= "Casein kinase II phosphorylation site"  
FT Modified-site 307..309  
FT /note= "Protein kinase C phosphorylation site"  
FT Modified-site 307..310  
FT /note= "Casein kinase II phosphorylation site"  
FT Modified-site 359..365  
FT /note= "Tyrosine kinase phosphorylation site"  
FT Modified-site 366..368  
FT Modified-site /note= "Protein kinase C phosphorylation site"  
FT Modified-site /note= "Casein kinase II phosphorylation site"  
FT Modified-site /note= "Casein kinase II phosphorylation site"  
FT Modified-site /note= "Protein kinase C phosphorylation site"  
FT Modified-site /note= "Protein kinase C phosphorylation site"  
FT Region 385..406  
FT /note= "Leucine zipper pattern"  
FT Modified-site 449..452  
FT /note= "Casein kinase II phosphorylation site"  
FT Modified-site 463..466  
FT /note= "Casein kinase II phosphorylation site"  
FT Modified-site 463..465  
FT /note= "Protein kinase C phosphorylation site"  
FT Modified-site 470..472  
FT /note= "Protein kinase C phosphorylation site"  
FT Modified-site 501..504  
FT /note= "Casein kinase II phosphorylation site"  
FT Modified-site 511..516  
FT /note= "N-myristoylation site"  
FT Domain 568..660  
FT /label= PDZ\_domain  
FT Modified-site 587..592  
FT /note= "N-myristoylation site"  
FT Modified-site 589..592  
FT /note= "N-glycosylation site"  
FT Modified-site 602..605  
FT /note= "Casein kinase II phosphorylation site"  
FT Modified-site 634..637  
FT /note= "Casein kinase II phosphorylation site"  
FT Modified-site 653..655  
FT /note= "Protein kinase C phosphorylation site"  
FT Modified-site 674..677  
FT /note= "Casein kinase II phosphorylation site"  
FT Domain 676..745  
FT /label= SH3\_domain  
FT Modified-site 714..719  
FT /note= "N-myristoylation site"  
FT Modified-site 725..727  
FT /note= "Protein kinase C phosphorylation site"  
FT Modified-site 725..728  
FT /note= "Casein kinase II phosphorylation site"  
FT Modified-site 733..738  
FT /note= "N-myristoylation site"  
FT Modified-site 737..740  
FT /note= "N-glycosylation site"  
FT Modified-site 759..761  
FT /note= "Protein kinase C phosphorylation site"  
FT Modified-site 760..763  
FT /note= "cAMP- and cGMP-dependent protein kinase  
phosphorylation site"  
FT Peptide 785..793  
FT /note= "Peroxisomal targeting signal"  
FT Modified-site 796..799  
FT /note= "Casein kinase II phosphorylation site"  
FT Modified-site 800..805  
FT /note= "N-myristoylation site"  
FT Domain 826..1004  
FT /label= Guanylate\_kinase\_domain  
FT Modified-site 842..844  
FT /note= "Protein kinase C phosphorylation site"  
FT Modified-site 860..863  
FT /note= "Casein kinase II phosphorylation site"  
FT Modified-site 868..870  
FT /note= "Protein kinase C phosphorylation site"  
FT Region 870..872  
FT /note= "RGD cell attachment sequence"  
FT Modified-site 893..896  
FT /note= "Casein kinase II phosphorylation site"  
FT Modified-site 926..929  
FT /note= "Casein kinase II phosphorylation site"  
FT Peptide 941..949

Query	Subject	Score	Length	DB	Accession	Description
1	MGELCRDPSALTALDEFTLWMMESHRRIRVRCICPSRLTPYLQAKYLQGLDEEVLHS	60	1004	1	MGELCRDPSALTALDEFTLWMMESHRRIRVRCICPSRLTPYLQAKYLQGLDEEVLHS	Modified-site
2	1	85.8%	Score 617.5	DB 22	Length 1004	/note="Peroxisomal targeting signal"
3	Best Local Similarity	90.4%	Pred. No. 4.8e-68			
4	Matches 123	Conservative 4	Mismatches 6	Indels 3	Gaps 2	
5	1	MGELCRDPSALTALDEFTLWMMESHRRIRVRCICPSRLTPYLQAKYLQGLDEEVLHS	60	1	MGELCRDPSALTALDEFTLWMMESHRRIRVRCICPSRLTPYLQAKYLQGLDEEVLHS	Modified-site
6	1	85.8%	Score 617.5	DB 22	Length 1004	/note="Casein kinase II phosphorylation site"
7	Best Local Similarity	90.4%	Pred. No. 4.8e-68			
8	Matches 123	Conservative 4	Mismatches 6	Indels 3	Gaps 2	
9	1	MGELCRDPSALTALDEFTLWMMESHRRIRVRCICPSRLTPYLQAKYLQGLDEEVLHS	60	1	MGELCRDPSALTALDEFTLWMMESHRRIRVRCICPSRLTPYLQAKYLQGLDEEVLHS	Modified-site
10	1	85.8%	Score 617.5	DB 22	Length 1004	/note="Casein kinase II phosphorylation site"
11	Best Local Similarity	90.4%	Pred. No. 4.8e-68			
12	Matches 123	Conservative 4	Mismatches 6	Indels 3	Gaps 2	
13	1	MGELCRDPSALTALDEFTLWMMESHRRIRVRCICPSRLTPYLQAKYLQGLDEEVLHS	60	1	MGELCRDPSALTALDEFTLWMMESHRRIRVRCICPSRLTPYLQAKYLQGLDEEVLHS	Modified-site
14	1	85.8%	Score 617.5	DB 22	Length 1004	/note="N-myristoylation site"
15	Best Local Similarity	90.4%	Pred. No. 4.8e-68			
16	Matches 123	Conservative 4	Mismatches 6	Indels 3	Gaps 2	
17	1	MGELCRDPSALTALDEFTLWMMESHRRIRVRCICPSRLTPYLQAKYLQGLDEEVLHS	60	1	MGELCRDPSALTALDEFTLWMMESHRRIRVRCICPSRLTPYLQAKYLQGLDEEVLHS	Modified-site
18	1	85.8%	Score 617.5	DB 22	Length 1004	/note="N-myristoylation site"
19	Best Local Similarity	90.4%	Pred. No. 4.8e-68			
20	Matches 123	Conservative 4	Mismatches 6	Indels 3	Gaps 2	
21	1	MGELCRDPSALTALDEFTLWMMESHRRIRVRCICPSRLTPYLQAKYLQGLDEEVLHS	60	1	MGELCRDPSALTALDEFTLWMMESHRRIRVRCICPSRLTPYLQAKYLQGLDEEVLHS	Modified-site
22	1	85.8%	Score 617.5	DB 22	Length 1004	/note="N-myristoylation site"
23	Best Local Similarity	90.4%	Pred. No. 4.8e-68			
24	Matches 123	Conservative 4	Mismatches 6	Indels 3	Gaps 2	
25	1	MGELCRDPSALTALDEFTLWMMESHRRIRVRCICPSRLTPYLQAKYLQGLDEEVLHS	60	1	MGELCRDPSALTALDEFTLWMMESHRRIRVRCICPSRLTPYLQAKYLQGLDEEVLHS	Modified-site
26	1	85.8%	Score 617.5	DB 22	Length 1004	/note="N-myristoylation site"
27	Best Local Similarity	90.4%	Pred. No. 4.8e-68			
28	Matches 123	Conservative 4	Mismatches 6	Indels 3	Gaps 2	
29	1	MGELCRDPSALTALDEFTLWMMESHRRIRVRCICPSRLTPYLQAKYLQGLDEEVLHS	60	1	MGELCRDPSALTALDEFTLWMMESHRRIRVRCICPSRLTPYLQAKYLQGLDEEVLHS	Modified-site
30	1	85.8%	Score 617.5	DB 22	Length 1004	/note="N-myristoylation site"
31	Best Local Similarity	90.4%	Pred. No. 4.8e-68			
32	Matches 123	Conservative 4	Mismatches 6	Indels 3	Gaps 2	
33	1	MGELCRDPSALTALDEFTLWMMESHRRIRVRCICPSRLTPYLQAKYLQGLDEEVLHS	60	1	MGELCRDPSALTALDEFTLWMMESHRRIRVRCICPSRLTPYLQAKYLQGLDEEVLHS	Modified-site
34	1	85.8%	Score 617.5	DB 22	Length 1004	/note="N-myristoylation site"
35	Best Local Similarity	90.4%	Pred. No. 4.8e-68			
36	Matches 123	Conservative 4	Mismatches 6	Indels 3	Gaps 2	
37	1	MGELCRDPSALTALDEFTLWMMESHRRIRVRCICPSRLTPYLQAKYLQGLDEEVLHS	60	1	MGELCRDPSALTALDEFTLWMMESHRRIRVRCICPSRLTPYLQAKYLQGLDEEVLHS	Modified-site
38	1	85.8%	Score 617.5	DB 22	Length 1004	/note="N-myristoylation site"
39	Best Local Similarity	90.4%	Pred. No. 4.8e-68			
40	Matches 123	Conservative 4	Mismatches 6	Indels 3	Gaps 2	
41	1	MGELCRDPSALTALDEFTLWMMESHRRIRVRCICPSRLTPYLQAKYLQGLDEEVLHS	60	1	MGELCRDPSALTALDEFTLWMMESHRRIRVRCICPSRLTPYLQAKYLQGLDEEVLHS	Modified-site
42	1	85.8%	Score 617.5	DB 22	Length 1004	/note="N-myristoylation site"
43	Best Local Similarity	90.4%	Pred. No. 4.8e-68			
44	Matches 123	Conservative 4	Mismatches 6	Indels 3	Gaps 2	
45	1	MGELCRDPSALTALDEFTLWMMESHRRIRVRCICPSRLTPYLQAKYLQGLDEEVLHS	60	1	MGELCRDPSALTALDEFTLWMMESHRRIRVRCICPSRLTPYLQAKYLQGLDEEVLHS	Modified-site
46	1	85.8%	Score 617.5	DB 22	Length 1004	/note="N-myristoylation site"
47	Best Local Similarity	90.4%	Pred. No. 4.8e-68			
48	Matches 123	Conservative 4	Mismatches 6	Indels 3	Gaps 2	
49	1	MGELCRDPSALTALDEFTLWMMESHRRIRVRCICPSRLTPYLQAKYLQGLDEEVLHS	60	1	MGELCRDPSALTALDEFTLWMMESHRRIRVRCICPSRLTPYLQAKYLQGLDEEVLHS	Modified-site
50	1	85.8%	Score 617.5	DB 22	Length 1004	/note="N-myristoylation site"
51	Best Local Similarity	90.4%	Pred. No. 4.8e-68			
52	Matches 123	Conservative 4	Mismatches 6	Indels 3	Gaps 2	
53	1	MGELCRDPSALTALDEFTLWMMESHRRIRVRCICPSRLTPYLQAKYLQGLDEEVLHS	60			

XX	Homo sapiens.	
OS		
FX	Key	Location/Qualifiers
FX	Domain	6..112
FT	Modified-site	/note= "CARD domain"
FT	Modified-site	7..9
FT	Modified-site	/note= "Protein kinase C phosphorylation site"
FT	Modified-site	7..10
FT	Modified-site	/note= "Casein kinase II phosphorylation site"
FT	Modified-site	100..102
FT	Modified-site	/note= "Protein kinase C phosphorylation site"
FT	Modified-site	100..103
FT	Modified-site	/note= "Casein kinase II phosphorylation site"
FT	Modified-site	105..107
FT	Modified-site	/note= "Protein kinase C phosphorylation site"
FT	Modified-site	106..109
FT	Modified-site	/note= "cAMP- and cGMP-dependent protein kinase phosphorylation site"
FT	Domain	130..431
FT	Modified-site	/note= "coiled coil domain"
FT	Modified-site	162..165
FT	Modified-site	/note= "Casein kinase II phosphorylation site"
FT	Modified-site	168..171
FT	Modified-site	/note= "Casein kinase II phosphorylation site"
FT	Modified-site	175..183
FT	Modified-site	/note= "Tyrosine kinase phosphorylation site"
FT	Modified-site	182..185
FT	Modified-site	/note= "Casein kinase II phosphorylation site"
FT	Modified-site	189..195
FT	Modified-site	/note= "Tyrosine kinase phosphorylation site"
FT	Modified-site	241..244
FT	Modified-site	/note= "N-glycosylation site"
FT	Modified-site	243..245
FT	Modified-site	/note= "Protein kinase C phosphorylation site"
FT	Modified-site	282..285
FT	Modified-site	/note= "Amidation site"
FT	Modified-site	286..289
FT	Modified-site	/note= "Casein kinase II phosphorylation site"
FT	Modified-site	290..292
FT	Modified-site	/note= "Protein kinase C phosphorylation site"
FT	Modified-site	378..381
FT	Modified-site	/note= "Casein kinase II phosphorylation site"
FT	Modified-site	429..432
FT	Modified-site	/note= "cAMP- and cGMP-dependent protein kinase phosphorylation site"
FT	Modified-site	459..461
FT	Modified-site	/note= "Protein kinase C phosphorylation site"
FT	Modified-site	471..474
FT	Modified-site	/note= "Casein kinase II phosphorylation site"
FT	Modified-site	472..475
FT	Modified-site	/note= "N-glycosylation site"
FT	Modified-site	476..479
FT	Modified-site	/note= "Casein kinase II phosphorylation site"
FT	Modified-site	508..510
FT	Modified-site	/note= "Protein kinase C phosphorylation site"
FT	Modified-site	510..513
FT	Modified-site	/note= "cAMP- and cGMP-dependent protein kinase phosphorylation site"
FT	Modified-site	558..560
FT	Modified-site	/note= "Protein kinase C phosphorylation site"
FT	Modified-site	578..581
FT	Modified-site	/note= "Casein kinase II phosphorylation site"
FT	Modified-site	584..587
FT	Modified-site	/note= "N-glycosylation site"
FT	Modified-site	587..592
FT	Modified-site	/note= "N-myristoylation site"
FT	Modified-site	634..637
FT	Modified-site	/note= "cAMP- and cGMP-dependent protein kinase phosphorylation site"

```

FT Domain 635..748 /note="PDZ domain"
FT Domain 635..1147 /note="MAGUK domain"
FT Modified-site 638..641 /note="MAGUK domain"
FT Modified-site 678..683 /note="Glycosaminoglycan attachment site"
FT Modified-site 687..689 /note="N-myristoylation site"
FT Modified-site 692..695 /note="Protein kinase C phosphorylation site"
FT Modified-site 698..703 /note="Casein kinase II phosphorylation site"
FT Modified-site 710..715 /note="N-myristoylation site"
FT Modified-site 725..728 /note="N-myristoylation site"
FT Modified-site 761..766 /note="Casein kinase II phosphorylation site"
FT Modified-site 764..767 /note="N-myristoylation site"
FT Modified-site 766..834 /note="Casein kinase II phosphorylation site"
FT Domain 776..779 /note="SH3 domain"
FT Modified-site 779..782 /note="N-glycosylation site"
FT Modified-site 787..789 /note="Casein kinase II phosphorylation site"
FT Modified-site 816..819 /note="Protein kinase C phosphorylation site"
FT Modified-site 823..828 /note="Casein kinase II phosphorylation site"
FT Modified-site 847..850 /note="N-myristoylation site"
FT Modified-site 853..858 /note="Casein kinase II phosphorylation site"
FT Modified-site 857..859 /note="N-myristoylation site"
FT Modified-site 872..875 /note="Protein kinase C phosphorylation site"
FT Modified-site 882..1147 /note="Casein kinase II phosphorylation site"
FT Domain 897..900 /note="Guanylate kinase (GUK) domain"
FT Modified-site 917..922 /note="Casein kinase II phosphorylation site"
FT Modified-site 926..929 /note="N-myristoylation site"
FT Modified-site 935..937 /note="Casein kinase II phosphorylation site"
FT Modified-site 1003..1006 /note="Protein kinase C phosphorylation site"
FT Modified-site 1010..1018 /note="Casein kinase II phosphorylation site"
FT Modified-site 1050..1055 /note="Tyrosine kinase phosphorylation site"
FT Modified-site 1088..1091 /note="N-myristoylation site"
FT Modified-site 1120..1123 /note="Casein kinase II phosphorylation site"
FT Modified-site 1120..1123 /note="Casein kinase II phosphorylation site"

```

```

XX 01-DEC-2000; 2000MO-US32716.
XX
XX 03-DEC-1999; 99US-0168780.
XX 18-FEB-2000; 2000US-0507533.
XX 25-FEB-2000; 2000US-0513904.
XX 10-OCT-2000; 2000US-0685791.

```

```

XX (MILL-) MILLENNIUM PHARM INC.
PA Bertin J;
XX WPI: 2001-367809/38.
XX N-PEDB; AAS05389.
DR
XX Novel caspase recruitment domain (CARD) proteins, CARD-9, CARD-10,
PT CARD-11, useful as targets for therapy, as immunogens, and in screening
PT and detection assays -
XX
PS Claim 9; Fig 14A-14C; 145pp; English.
XX
XX The present sequence represents novel human caspase recruitment
XX domain, CARD-11. The polynucleotide encoding this sequence was
XX isolated from a human T-cell cDNA library. Also described are
XX novel human sequences for CARD-9 and CARD-10 (AAU01205, AAU01206)
XX and rat CARD-9 (AAU01204). CARD-9, CARD-10 and CARD-11 interact
XX with Bcl-10 which is thought to activate nuclear factor (NF)-kappaB
XX and apoptosis. The sequences of the invention can be used for
XX treating a disorder associated with abnormal levels of apoptosis by
XX modulating the expression or activity of CARD-9, CARD-10, or CARD-11.
XX They can be used for the treatment of hyperproliferative disorders
XX (e.g. cancer), autoimmune disorders (e.g. systemic lupus erythematosus),
XX neurological disorders (e.g. Alzheimer's disease), inflammatory disorders
XX (e.g. Crohn's disease), and viral infection (e.g. HIV). The CARD
XX polypeptide, polynucleotide and an antibody which selectively binds to
XX CARD can be used in screening and detection assays (e.g. chromosomal
XX mapping, tissue typing), predictive medicine (prognostic assays),
XX monitoring clinical trials, and therapy (treatment and prophylaxis). The
XX CARD polypeptide may be used to screen for drugs that bind to and/or
XX modulate it. CARD sequences are potential targets for regulating
XX inflammation, cancer, NF-kappaB signaling, stress-related response and
XX apoptosis in human disease. A host cell containing a polynucleotide
XX encoding CARD can be used to create transgenic animals.
XX
SQ Sequence 1147 AA;
Query Match 37.6%; Score 270.5; DB 22; Length 1147;
Best Local Similarity 48.6%; Pred. No. 2.9e-24;
Matches 54; Conservative 23; Mismatches 33; Indels 1; Gaps 1;
QY 15 DEETMEMESHRRIRVCIQPSLTPTLRQAKLQCDREYVHSPRLTMSARAGTLL 74
DB 11 EEDALMEVENCNRMLSRYINPAKLTPLRQCKVIDQDEDEVANAPVLPKINRAGRL 70
QY 75 DLKTRGNKAIAFLESILKFNPDVYTLVGLQDPVPSNFGSSDPDGL 125
DB 71 DILHTKGGRGVVFLSLEFYFPELVKLVTKKPTRRPSTIVZEG-HEGL 120

```

## RESULT 5

AAG79555 standard; Protein, 1147 AA.

AAG79555;

09-DEC-2002 (first entry)

Human CARD-11.

XX

XX Rac, human; caspase recruitment domain; CARD-9; CARD-10;

XX CARD-11; apoptosis; inflammation; cell growth; cell death;

XX lymphocyte activation; cancer; melanoma; autoimmune disease;

XX arthritis; neurological disorder; Alzheimer's disease.

XX

OS Homo sapiens.

XX

XX Key

XX Location/Qualifiers

XX 6..112

XX /label= CARD\_domain

XX 130..431

XX Domain

	FT	Domain	/label= Coiled-coil_domain 635..748
	PT	Domain	/label= PDZ_domain
	FT	Domain	766..834
	FT	Domain	/label= SH3_domain
	PT	Domain	882..1147
	FT	Domain	/label= Guanylate_kinase_domain
	PT	Domain	635..1147
	FT	Domain	/label= MAGUK_domain
XN	WO200270652-A2.		
XX	12-SEP-2002.		
PB			
PF	28-FEB-2002; 2002WQ-US06147.		
XX			
PR	02-MAR-2001; 2001US-0798412.		
XX	(MILL-) MILLENNIUM PHARM INC.		
PA	Bertin J;		
PI			
DH	WPJ; 2002-698749/75.		
DR	N-PDB; ABAO0334.		
XX			
PS	This sequence represents human caspase recruitment domain (CARD)-11. CARD proteins play roles in apoptotic and inflammatory signalling pathways. CARD-9, -10 and -11 participate in the network of interactions that modulate caspase activity. They are thought to be useful as modulating agents for regulating a variety of cellular processes including cell growth and cell death. CARD proteins and nucleic acids are useful for treating a disorder associated with inappropriate apoptosis or lymphocyte activation or for diagnosing subjects having or that are at risk of developing a disorder associated with aberrant CARD-9, CARD-10 or CARD-11 expression or activity, such as cancer e.g. melanoma, autoimmune disorders e.g. arthritis, or neurological disorders e.g. Alzheimer's disease.		
CC	Sequence 1147 AA;		
SQ			
	Query Match 37.6%; Score 270.5; DB 23; Length 1147; Best Local Similarity 48.6%; Pred. No. 2.9e+24; Matches 54; Conservative 23; Mismatches 33; Indels 1; Gaps 14		
OY	15 DEETLWMEMSHRRIRIVRCIPSRLLTPYLROAKVLCQDEESVLASPLRLTNSAMRAGHL 74 :: Db 11 EEDALMEWEVCNRMLSLRYINPAAKLTFLTRCKCVIDDEDDEVLANAPLPKINKRAGL 70 ::		
OY	75 DLATRGNGAIAFLESKFTHNPDYITLVTLGQPVDVSFNFSGESSDPDL 125 :: Dd 71 DILHTKGGRGVVFLESLEFYYPKLKVTKGPFRFRSTIVEBG-HEGL 120 ::		
RESULT 6			
ABG76061	ID ABG76061 standard; protein; 1247 AA.		
AC	ABG76061;		
XX			
DT	09-MAY-2003 (first entry)		
DE	Human caspase recruitment domain containing protein, CARD-11X.		
KW	CARD; caspase recruitment domain; apoptosis; cell adhesion; inflammation; cytokine receptor signaling; cancer; glioma; carcinoma; adenocarcinoma; CARD-containing polypeptide associated disorder; sarcoma; melanoma; hamartoma; leukemia; lymphoma; keratinocyte hyperplasia; neoplasia;		

XX	Keloid; benign prostatic hypertrophy; inflammatory hyperplasia; fibrosis;
XW	restenosis; allergy; arthritis; lupus; Sjogren's syndrome; sepsis; human;
KM	Crohn's disease; ulcerative colitis; graft versus host disease; stroke;
KW	abnormal cell death disease; myocardial infarction; heart failure;
KV	neurodegenerative disease; Parkinson's disease; Alzheimer's disease; HIV;
KX	CARD-11x; caspase activator; caspase inhibitor.
OS	Homo sapiens.
XX	
FH	Key Location/Qualifiers
FT	Domain 12..103
FT	/label= CARD
FT	/note= "Caspase recruitment domain. Specifically
FT	claimed in claim 12"
FT	Domain 368..1344
FT	/label= ERN domain
FT	/note= "Erin/radixin/moesin domain. Specifically
FT	claimed in claim 12"
PN	Domain 2175..2514
XX	/label= PDZ domain
PD	/note= "Post synaptic density disc-large zo-1 domain.
XX	Specifically claimed in claim 12"
PF	US2002164703-A1.
XX	
XX	07-NOV-2002.
XX	
XX	19-DEC-2001; 2001US-0032159.
PR	21-DEC-2000; 2000US-257457P.
PA	(PATL/) PAWLOWSKI K.
PA	(REED/) REED J C.
PA	(GODZ/) GODZIK A.
PI	Pawlowski K, Reed JC, Godzik A;
DR	WPI; 2003-288137/28.
DR	N-PDB; ABX11430.
PT	New isolated CARD-containing nucleic acids, useful for the diagnosis
PT	and treatment of disorders with aberrant expression or activity of the
PT	CARD-containing polypeptide, such as cancer, stroke, arthritis, heart
PS	failure and AIDS -
XX	Claim 11; Fig 2; 34pp; English.
CC	The invention relates to an isolated nucleic acid molecule encoding a
CC	caspase recruitment domain (CARD) containing polypeptide. CARD containing
CC	polypeptides are involved in apoptosis (as caspase activators and caspase
CC	inhibitors), cell adhesion, inflammation and cytokine receptor
CC	signalling. The methods and compositions of the present invention are
CC	useful for the diagnosis and treatment of disorders associated with the
CC	aberrant expression or activity of the CARD containing polypeptide such
CC	as cancer, glioma, carcinoma, adenocarcinoma, sarcoma, melanoma, keloid,
CC	hamartoma, leukemia, lymphoma, keratinocyte hyperplasia, neoplasia,
CC	benign prostatic hypertrophy, inflammatory hyperplasia, fibrosis,
CC	restenosis, allergies, inflammatory diseases such as arthritis, lupus,
CC	Sjogren's syndrome, Crohn's disease, ulcerative colitis, graft versus
CC	host disease, sepsis, abnormal cell death diseases such as stroke
CC	myocardial infarction, heart failure, neurodegenerative diseases like
CC	Parkinson's disease and Alzheimer's disease, and HIV infection. The
CC	present sequence represents the amino acid sequence of the caspase
CC	recruitment domain containing protein, CARD-11x.
SO	Sequence 1247 AA;
Query Match	37.6%; Score 270.5; DB 24; Length 1247;
Best Local Similarity	48.6%; Pred. No. 3.3e-24;
Matches 54; Conservative 23; Mismatches 33; Indels 1; Gaps 1;	
15 DEBTLMAMESHRRIYRCICPSRLTYLYIAQAKLCGLDSEEVLIHSPLTNMSNRAGHL 74	

Db 11 EEDALWENECNRHMLSRYPINPAKLTPYLRQCKVIDEODEVLNAPMLPSKINRAGRLL 70  
QY 75 DLKTRGKGAIAFLESLKFNHPDYVTLVGTGQPDVDFNFGESDFRGL 125  
Db 71 DILHTKGQRYVFLFSLFEPYLPETKYLGKPTFRFSTIVEEG-HRGL 120

RESULT 7  
AAU01206  
ID AAU01206 standard; Protein; 1032 AA.  
XX AAU01206;  
XX  
DT 12-SEP-2001 (first entry)  
XX Human caspase recruitment domain, CARD-10 polypeptide.  
XX Human; caspase recruitment domain; CARD-10; Bcl-10; NF-kappaB;  
KM apoptosis; hyperproliferative disorder; autoimmune; neurological;  
KM inflammatory disorder; viral infection; stress-related response.  
XX Homo sapiens.  
XX

Key Location/Qualifiers  
FH Modified-site 15..20 /note= "N-myristoylation site"  
FT Modified-site 18..21 /note= "Casein kinase II phosphorylation site"  
FT Domain 23..123 /note= "CARD domain"  
FT Modified-site 68..70 /note= "Protein kinase C phosphorylation site"  
FT Modified-site 76..79 /note= "N-glycosylation site"  
FT Modified-site 78..80 /note= "Protein kinase C phosphorylation site"  
FT Modified-site 88..91 /note= "Amidation site"  
FT Modified-site 112..115 /note= "Casein kinase II phosphorylation site"  
FT Modified-site 113..118 /note= "N-myristoylation site"  
FT Domain 147..457 /note= "Coiled coil domain"  
FT Modified-site 201..207 /note= "Tyrosine kinase phosphorylation site"  
FT Region 230..251 /note= "Leucine zipper homology region"  
FT Modified-site 242..245 /note= "Casein kinase II phosphorylation site"  
FT Modified-site 293..295 /note= "Protein kinase C phosphorylation site"  
FT Modified-site 293..296 /note= "Casein kinase II phosphorylation site"  
FT Modified-site 309..314 /note= "N-myristoylation site"  
FT Modified-site 313..315 /note= "Protein kinase C phosphorylation site"  
FT Modified-site 331..334 /note= "Casein kinase II phosphorylation site"  
FT Domain 366..398 /note= "Tropomyosin domain"  
FT Modified-site 412..415 /note= "Casein kinase II phosphorylation site"  
FT Region 426..447 /note= "Leucine zipper homology region"  
FT Modified-site 438..441 /note= "Casein kinase II phosphorylation site"  
FT Domain 457..1032 /note= "N-myristoylation site"  
FT Modified-site 472..475 /note= "N-glycosylation site"  
FT Modified-site 478..481

FT Modified-site /note= "Casein kinase II phosphorylation site"  
FT /note= "N-myristoylation site"  
FT Modified-site 510..513 /note= "Casein kinase II phosphorylation site"  
FT Modified-site 512..514 /note= "Protein kinase C phosphorylation site"  
FT Modified-site 549..552 /note= "Casein kinase II phosphorylation site"  
FT Modified-site 558..560 /note= "Protein kinase C phosphorylation site"  
FT Modified-site 565..570 /note= "N-myristoylation site"  
FT Modified-site 570..573 /note= "Casein kinase II phosphorylation site"  
FT Modified-site 595..598 /note= "N-glycosylation site"  
FT Modified-site 603..605 /note= "Protein kinase C phosphorylation site"  
FT Modified-site 638..641 /note= "Glycosaminoglycan attachment site"  
FT Modified-site 642..644 /note= "Protein kinase C phosphorylation site"  
FT Modified-site 656..661 /note= "N-myristoylation site"  
FT Modified-site 681..684 /note= "Casein kinase II phosphorylation site"  
FT Modified-site 690..693 /note= "Casein kinase II phosphorylation site"  
FT Domain 704..772 /note= "SH3 domain"  
FT Modified-site 712..715 /note= "N-glycosylation site"  
FT Modified-site 714..717 /note= "Casein kinase II phosphorylation site"  
FT Modified-site 733..739 /note= "Tyrosine kinase phosphorylation site"  
FT Modified-site 748..751 /note= "Casein kinase II phosphorylation site"  
FT Modified-site 754..756 /note= "Protein kinase C phosphorylation site"  
FT Modified-site 754..757 /note= "Casein kinase II phosphorylation site"  
FT Modified-site 761..766 /note= "N-myristoylation site"  
FT Modified-site 782..784 /note= "Protein kinase C phosphorylation site"  
FT Modified-site 809..814 /note= "N-myristoylation site"  
FT Domain 830..1032 /note= "Guanylate kinase (GUK) domain"  
FT Modified-site 830..832 /note= "Protein kinase C phosphorylation site"  
FT Modified-site 868..870 /note= "Protein kinase C phosphorylation site"  
FT Modified-site 869..872 /note= "Casein kinase II phosphorylation site"  
FT Modified-site 882..885 /note= "Casein kinase II phosphorylation site"  
FT Modified-site 893..898 /note= "Casein kinase II phosphorylation site"  
FT Modified-site 915..918 /note= "Amidation site"  
FT Modified-site 947..949 /note= "Protein kinase C phosphorylation site"  
FT Modified-site 981..986 /note= "N-myristoylation site"  
FT Modified-site 1021..1026 /note= "N-myristoylation site"  
FT Modified-site 1022..1024 /note= "Protein kinase C phosphorylation site"  
FT Modified-site 1028..1031 /note= "Casein kinase II phosphorylation site"

[illegible]

KW		Ret : human; caspase recruitment domain; CARD-9; CARD-10;
KM	CARD-11; apoptosis; inflammation; cell growth; cell death;	
KV	Lymphocyte activation; cancer; melanoma; autoimmune disease;	
KW	arthritis; neurological disorder; Alzheimer's disease.	
XX	Homo sapiens.	
OS		
FH	Key	Location/Qualifiers
FT	Domain	23..123
FT	Domain	/label= CARD_domain
FT	Domain	147..457
FT	Domain	/label= Colled-coil_domain
FT	Domain	704..772
FT	Domain	/label= SH3_domain
FT	Domain	830..1032
FT	Domain	/label= Guanylate_kinase_domain
FT	Domain	366..398
FT	Domain	/label= Tropomyosin_domain
FT	Domain	457..1032
FT	Domain	/label= MAGUK_domain
XX		
PN	WC0200270652-A2.	
PD	12-SEP-2002.	
XX		
PF	28-FEB-2002; 2002MCO-USO6147.	
PR	02-MAR-2001; 2001US-0798412.	
PA	(MTL-) MILLENNIUM PHARM INC.	
XK	Bertin J;	
DR	WPI ; 2002-698749/75.	
N-	PSDB; ABA00333.	
XX		
PT	CARD-9, CARD-10 or CARD-11 polypeptides and polymucleotides, useful for treating disorders associated with inappropriate apoptosis or lymphocyte activation, e.g. cancer -	
PS	Claim 14; Fig 10; 151pp; English.	
CC	This sequence represents human caspase recruitment domain (CARD)-10. CARD proteins play roles in apoptotic and inflammatory signalling pathways. CARD-9, -10 and -11 participate in the network of interactions that modulate caspase activity. They are thought to be useful as modulating agents for regulating a variety of cellular processes including cell growth and cell death. CARD proteins and nuclear acids are useful for treating a disorder associated with inappropriate apoptosis or lymphocyte activation or for diagnosing subjects having or that are at risk of developing a disorder associated with aberrant CARD-9, CARD-10 or CARD-11 expression or activity, such as cancer e.g. melanoma, autoimmune disorders e.g. arthritis, or neurological disorders e.g. Alzheimer's disease.	
SQ	Sequence 1032 AA:	
Dz	Query Match 34.2%; Score 246; DB 23; Length 1032; Best Local Similarity 48.9%; Pred. No. 3e+21; Matches 46; Conservative 21; Mismatches 27; Indels 0; Gaps 0;	
Oy	15 DEETLTMMESHRRIVRCICPSSRLTPYLNQAAYLCOLDEEVLHSPLTNSAMRAGHL 74 :: ::  :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: ::  Db 23 EEDTLMERIGSVHRALRNAPAKLPYYLAQCVIDBQBEEVLYSTVFPCRVVRTRLM 82 ::	
Oy	75 DLKTGRGNGLAFLBSLKFAHPNDVTITLVGLOP 108  ::  Db 83 DILRGCKRGVEAFLEALEFYEPHEHTLGTGOBP 116  ::	
RESULT 9		
AU073247		
ID AU073247 standard; Protein; 1032 AA.		

```

XX AC AAU73247;
XX DT 12-MAR-2002 (first entry)
XX DE Human plakoglobin interacting protein #3.
XX KM Human; plakoglobin; cytosolic; osteopathic; dermatological; cardiac;
XX KM plakoglobin related disease; skin carcinoma; acantholytic disease;
XX KM basal cell carcinoma; squamous cell carcinoma; Naxos disease;
XX KM extramammary Paget's disease; heart disease; skin blistering;
XX KM subcorneal acantholysis; Grover's disease; Halley-Halley's disease;
XX KM Darier's disease; ectodermal dysplasia; skin fragility syndrome.
XX OS Homo sapiens.
XX PN MO200185933-A2.
XX PD 15-NOV-2001.
XX PF 02-MAY-2001; 2001MO-EP04872.
XX PR 09-MAY-2000; 2000EP-0201668.
XX PA (VLAAs-) VLAAMS INTERUNIVERSITAIR INST BIOTECHNOG.
XX PI Van Roy F, Bonne S, Vanlandschoot A;
XX DR WPI: 2002-063246/08.
XX DR N-PSDB; AAS98203.
XX PT New polypeptide, useful for treating skin carcinoma or acantholytic
XX PT disease such as Grover's and Darier's disease, comprises a protein
XX PT interacting with human plakoglobin and involved in transduction of
XX PT plakoglobin related signal to nucleus -
XX PS Claim 1; Figure 3; 98pp; English.
XX CC The invention relates to an isolated plakoglobin interacting polypeptide
XX CC (1). (1) is useful as a medicament and in the manufacture of a
XX CC medicament for treating plakoglobin related diseases, such as skin
XX CC carcinoma or an acantholytic disease, and to screen compounds that
XX CC interfere with the interaction of the polypeptide with plakoglobin
XX CC The plakoglobin related diseases include basal cell carcinoma, squamous
XX CC cell carcinoma, extramammary Paget's disease, Naxos disease, heart
XX CC disease, skin blistering and acantholytic diseases such as subcorneal
XX CC acantholysis, Grover's disease, Halley-Halley's disease or Darier's
XX CC disease, and ectodermal dysplasia/skin fragility syndrome. AAU73245-
XX CC AAU73254 represent novel human plakoglobin interacting protein
XX CC amino acid sequences and related sequences of the invention.
XX SQ Sequence 1032 AA;
XX
XX Query Match 34.2%; Score 246; DB 23; Length 1032;
XX Best Local Similarity 48.9%; Pred. No 3e-21;
XX Matches 46; Conservative 21; Mismatches 27; Indels 0; Gaps 0;
XX
XX QY 15 DEFTLWMESHRRHRYVRCIPSRLLPYLQAAVLCQLEEEVYHSRPLTNSMRAGHLL 74
XX DB 23 EEDLWLRIGVGRRLARALNPAKLTPLYLQRCVVIDEODEEVLSTYRPFRCVNRGRML 82
XX QY 75 DLKTRGKGAIAFLBSLKFHPDVTYLVGTLP 108
XX DB 83 DILRCRGRGYEAFLEALFEYFPEHFTLTGQEP 116
XX
XX RESULT 10
XX AAU01204
XX ID AAU01204 standard; Protein; 536 AA.
XX AC AAU01204;
XX DT 12-SEP-2001 (first entry)

```

```

XX DE Rat caspase recruitment domain, CARD-9 polypeptide.
XX KM Rat; caspase recruitment domain; CARD-9; Bcl-10; NF-kappaB;
XX KM apoptosis; hyperproliferative disorder; autoimmune; neurological;
XX KM inflammatory disorder; viral infection; stress-related response.
XX OS Rattus sp.
XX PN Key
XX PD Modified-site
XX FT Domain
XX FT /note= "CARD domain"
XX FT /note= "Casein kinase II phosphorylation site"
XX FT 7..98
XX FT /note= "CARD domain"
XX FT /note= "Casein kinase II phosphorylation site"
XX FT 12..15
XX FT /note= "Casein kinase II phosphorylation site"
XX FT 16..18
XX FT /note= "Protein kinase C phosphorylation site"
XX FT 23..26
XX FT /note= "Casein kinase II phosphorylation site"
XX FT 92..95
XX FT /note= "CAMP- and CGMP-dependent protein kinase
XX FT phosphorylation site"
XX FT 95..97
XX FT /note= "Protein kinase C phosphorylation site"
XX FT 95..98
XX FT /note= "Casein kinase II phosphorylation site"
XX FT 138..140
XX FT /note= "Protein kinase C phosphorylation site"
XX FT 138..141
XX FT /note= "Casein kinase II phosphorylation site"
XX FT 140..416
XX FT /note= "Coiled coil domain"
XX FT 171..174
XX FT /note= "Casein kinase II phosphorylation site"
XX FT 176..183
XX FT /note= "Tyrosine kinase phosphorylation site"
XX FT 197..213
XX FT /note= "Indole-3-glycerol phosphate synthase
XX FT homology region"
XX FT 228..231
XX FT /note= "CAMP- and CGMP-dependent protein kinase
XX FT phosphorylation site"
XX FT 231..233
XX FT /note= "Protein kinase C phosphorylation site"
XX FT 267..270
XX FT /note= "Casein kinase II phosphorylation site"
XX FT 285..338
XX FT /note= "Cysteine rich repeat homology region"
XX FT 303..305
XX FT /note= "Protein kinase C phosphorylation site"
XX FT 362..364
XX FT /note= "Protein kinase C phosphorylation site"
XX FT 362..365
XX FT /note= "Casein kinase II phosphorylation site"
XX FT 374..377
XX FT /note= "Casein kinase II phosphorylation site"
XX FT 425..428
XX FT /note= "Casein kinase II phosphorylation site"
XX FT 431..433
XX FT /note= "Protein kinase C phosphorylation site"
XX FT 451..453
XX FT /note= "Protein kinase C phosphorylation site"
XX FT 483..486
XX FT /note= "Casein kinase II phosphorylation site"
XX FT 514..516
XX FT /note= "Protein kinase C phosphorylation site"
XX FT 523..528
XX FT /note= "N-myristoylation site"
XX FT 524..527
XX FT /note= "N-glycosylation site"
XX FT 526..529
XX FT /note= "Casein kinase II phosphorylation site"

```





XX 26-JUN-2001 (first entry)  
 XX Human protein sequence SEQ ID NO:18328.  
 DE Human; primer; detection; diagnosis; antisense therapy; gene therapy.  
 XX Homo sapiens.  
 OS Homo sapiens.  
 XX EPI074617-A2.  
 XX 07-FEB-2001.  
 XX 28-JUL-2000; 2000EP-0116126.  
 XX 29-JUL-1999; 99JP-0248036.  
 XX 27-AUG-1999; 99JP-0300253.  
 XX 11-JAN-2000; 2000JP-0118776.  
 XX 02-MAY-2000; 2000JP-0183767.  
 XX 09-JUN-2000; 2000JP-0241899.  
 XX (HELI-) HELIX RES INST.  
 PA Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;  
 PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;  
 DR WPI; 2001-318749/34.  
 XX Primer sets for synthesizing polynucleotides, particularly the 5602  
 PT full-length CDNA defined in the specification, and for the detection  
 PT and/or diagnosis of the abnormality of the proteins encoded by the  
 PT full-length CDNA.  
 PS Claim 8; SEQ ID 18328; 2537PP + CD ROM; English.  
 XX The present invention describes primer sets for synthesizing 5602  
 CC full-length CDNA defined in the specification. Where a primer set  
 CC comprises: (a) an oligo-dT primer and an oligonucleotide complementary  
 CC to the complementary strand of a polynucleotide which comprises one of  
 CC the 5602 nucleotide sequences defined in the specification, where the  
 CC oligonucleotide comprises at least 15 nucleotides; or (b) a combination  
 CC of an oligonucleotide comprising a sequence complementary to the  
 CC complementary strand of a polynucleotide which comprises a 5'-end  
 CC sequence and an oligonucleotide comprising a sequence complementary to a  
 CC polynucleotide which comprises a 3'-end sequence, where the  
 CC oligonucleotide comprises at least 15 nucleotides and the combination of  
 CC the 5'-end sequence/3'-end sequence is selected from those defined in  
 CC the specification. The primer sets can be used in antisense therapy and  
 CC in gene therapy. The primers are useful for synthesizing polynucleotides,  
 CC particularly full-length CDNA. The primers are also useful for the  
 CC detection and/or diagnosis of the abnormality of the proteins encoded by  
 CC the full-length CDNA. The primers allow obtaining of the full-length  
 CC CDNA easily without any specialised methods. AAH03166 to AAH13628 and  
 CC AAH1633 to AAH18742 represent human CDNA sequences; AAB92446 to  
 CC AAB95893 represent human amino acid sequences; and AAH13629 to AAH13632  
 CC represent oligonucleotides, all of which are used in the exemplification  
 CC of the present invention.  
 XX Sequence 366 AA;  
 SO  
 Query Match 29.9%; Score 215.5; DB 22; Length 366;  
 Best Local Similarity 44.4%; Pred. No. 4.5e-18;  
 Matches 48; Conservative 17; Mismatches 40; Indels 3; Gaps 1;  
 QY 15 DEETLWEMSHHRRIVRCICPSRLIPYRQAVYLCQDDEEVLHSPRTNSAMRAGHLL 74  
 DB 6 NDECKNVLEGFVTLTVSDIPSRITPYRJQCKVLPNDDSEQLSDPNVIRKRGVALL 65  
 QY 75 DLKTRGKNGAIAFLSLKFNHNDVYTVGLQPDVDF--NFSGES 119  
 DB 66 DILORTGHKGYVAFLESLLELYPOLYKRYGKEPAVFSMTIDASGES 113

RESULT 13  
 AAG76060  
 ID AAG76060 standard; protein; 366 AA.  
 XX  
 AC AAG76060;  
 XX  
 XX 09-MAY-2003 (first entry)  
 XX  
 DE Caspase recruitment domain containing protein, CARD-10X.  
 XX  
 XX CARD; caspase recruitment domain; apoptosis; cell adhesion; inflammation;  
 XX cytokine receptor signalling; cancer; glioma; carcinoma; adenocarcinoma;  
 XX CARD-containing polypeptide associated disorder; sarcoma; melanoma;  
 XX hamartoma; leukaemia; lymphoma; keratinocyte hyperplasia; neoplasia;  
 XX keloid; benign prostatic hypertrophy; inflammatory hyperplasia; fibrosis;  
 XX restenosis; allergy; arthritis; lupus; Sjogren's syndrome; sepsis;  
 XX Crohn's disease; ulcerative colitis; graft versus host disease; stroke;  
 XX abnormal cell death disease; myocardial infarction; heart failure;  
 XX neurodegenerative disease; Parkinson's disease; Alzheimer's disease; HIV;  
 XX CARD-10X; caspase activator; caspase inhibitor.  
 OS Unidentified.  
 XX  
 XX Key Location/Qualifiers  
 FH 12..83  
 FT Domain /label= CARD  
 FT /note= "Caspase recruitment domain. Specifically  
 FT 154..297  
 FT /label= Filament domain  
 FT /note= "Specifically claimed in claim 12"  
 XX  
 XX US2002164703-A1.  
 XX 07-NOV-2002.  
 XX 19-DEC-2001; 2001US-0032159.  
 XX PF 21-DEC-2000; 2000US-257457P.  
 XX  
 PA (PAWL/) PAWLOWSKI K.  
 PA (REED/) REED J C.  
 PA (GODZ/) GODZIK A.  
 PI Pawlowski K, Reed JC, Godzik A;  
 DR WPI; 2003-288137/28.  
 DR N-PSDB; ABX11432.  
 XX  
 XX New isolated CARD-containing nucleic acids, useful for the diagnosis  
 PT and treatment of disorders with aberrant expression or activity of the  
 PT CARD-containing polypeptide, such as cancer, stroke, arthritis, heart  
 PT failure and AIDS.  
 XX  
 XX Claim 13; Fig 1; 34pp; English.  
 PS  
 XX The invention relates to an isolated nucleic acid molecule encoding a  
 CC caspase recruitment domain (CARD) containing polypeptide. CARD containing  
 CC polypeptides are involved in apoptosis (as caspase activators and caspase  
 CC inhibitors), cell adhesion, inflammation and cytokine receptor  
 CC signalling. The methods and compositions of the present invention are  
 CC useful for the diagnosis and treatment of disorders associated with the  
 CC aberrant expression or activity of the CARD containing polypeptide such  
 CC as cancer, glioma, carcinoma, adenocarcinoma, sarcoma, melanoma, keloid,  
 CC hamartoma, leukaemia, lymphoma, keratinocyte hyperplasia, fibrosis,  
 CC benign prostatic hypertrophy, inflammatory hyperplasia, neoplasia,  
 CC restenosis, allergy, arthritis, lupus, Sjogren's syndrome, sepsis,  
 CC Crohn's disease, ulcerative colitis, graft versus host disease, stroke,  
 CC abnormal cell death disease, myocardial infarction, heart failure,  
 CC neurodegenerative disease, Parkinson's disease, and HIV infection. The  
 CC present sequence represents the amino acid sequence of the caspase  
 CC recruitment domain containing protein, CARD-10X.



CC neurological disorders e.g. Alzheimer's disease

Search completed: February 18, 2004, 04:44:28  
Job time : 77 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using SW model

Run on: February 18, 2004, 04:37.37 ; Search time 36 Seconds  
(Without alignments)  
371.318 Million cell updates/sec

Title: US-10-032-159A-16

Perfect score: 720  
Sequence: 1 MSLCRDRSALTALDEETLM.....SDPDGIAGTSRNLRLVTPX 139

Scoring table: BLOSUM62  
Gapop 10.0, Gapext 0.5

Searched: 283308 segs, 96168682 residues

Total number of hits satisfying chosen parameters: 283308

Minimum DB seg length: 0  
Maximum DB seg length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database :  
1: PIR 76:\*  
2: PIR1:\*  
3: PIR2:\*  
4: PIR3:\*  
5: PIR4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	81	11.2	579	2 F69157	exonuclease ABC c
2	80	11.1	2114	2 E96505	hypothetical prote
3	76.5	10.6	532	2 H83493	hypothetical prote
4	75	10.4	271	1 A89124	probable phospho
5	74	10.3	210	2 S5671	hypothetical prote
6	73.5	10.2	315	2 T26673	hypothetical prote
7	73.5	10.2	1113	2 T14260	period protein Per
8	73	10.1	265	2 H97280	hydroxyethylthiaz
9	72.5	10.1	1318	2 T21266	hypothetical prote
10	72.5	10.1	1383	2 S59644	sister chromatid c
11	72	10.0	229	2 I37468	house keeping gene
12	72	10.0	1234	2 B36186	I factor protein 2
13	71.5	9.9	660	2 T45569	receptor protein k
14	71	9.9	213	2 A90269	conserved hypothet
15	71	9.9	265	2 T33012	hypothetical prote
16	71	9.9	428	2 A40241	probable coenzyme
17	70.5	9.8	263	2 A30227	hypothetical prote
18	70.5	9.8	329	2 H83604	hypothetical prote
19	70.5	9.8	356	2 S13221	GTP-binding regula
20	70.5	9.8	573	2 S69181	protein disulfide-
21	70.5	9.8	2376	2 S48405	probable membrane
22	70	9.7	559	2 AB3735	beta-(1->2)glucan
23	69.5	9.7	349	2 AF2130	iron(III) diclitrat
24	69.5	9.7	522	2 D72349	conserved hypothet
25	69.5	9.7	1452	1 S17670	protein-tyrosine-p
26	69	9.6	496	2 E83849	spore germination
27	69	9.6	967	2 A30325	membrane alanyl am
28	68.5	9.5	532	2 A65138	hypothetical 60.3
29	68.5	9.5	532	2 A91162	probable 2-compone

30	68.5	9.5	532	2 B86008	probable 2-compone
31	68.5	9.5	557	2 A12239	hypothetical prote
32	68.5	9.5	1115	2 T13955	period protein Per
33	68	9.4	603	2 S34130	serine/threonine-s
34	67.5	9.4	270	2 T50952	transposase relate
35	67.5	9.4	460	2 E86922	probable FAD-lyase
36	67.5	9.4	521	2 AB0269	anthranilate synth
37	67.5	9.4	872	2 E71852	valine-tRNA ligase
38	67.5	9.4	882	1 L1HUC2	cadherin 1 precurs
39	67.5	9.4	1420	2 B57062	SRS9 protein - yea
40	67	9.3	404	2 F81724	conserved hypothet
41	67	9.3	493	2 T04961	hypothetical prote
42	67	9.3	728	2 S57142	hypothetical prote
43	67	9.3	867	2 H70411	alanine-tRNA ligas
44	67	9.3	1012	2 B90389	conserved hypothet
45	67	9.3	1446	2 T13018	hypothetical prote

## ALIGNMENTS

RESULT 1  
F69157  
exonuclease ABC chain C - Methanobacterium thermoautotrophicum (strain Delta H)  
C/Species: Methanobacterium thermoautotrophicum  
C/Date: 05-Dec-1997 #sequence\_revision 05-Dec-1997 #text\_change 20-Apr-2000  
C/Accession: F69157  
R/Smith, D.R.; Doucette-Stamm, L.A.; Deloughery, C.; Lee, H.; Dubois, J.; Aldredge, T.;  
Kl. S.; Church, G.M.; Daniels, C.J.; Mao, J.; Rice, P.; Noelling, J.; Reeve, J.N.  
J. Bacteriol. 179, 7135-7155, 1997  
A/Title: Complete genome sequence of Methanobacterium thermoautotrophicum Delta H: func  
A/Reference number: A69000; MUID:98037514; PMID:9371463  
A/Accession: F69157  
A/Status: preliminary; nucleic acid sequence not shown; translation not shown  
A/Molecule type: DNA  
A/Residues: 1-579 <MT>  
A/Cross-references: GB:A6000828; GB:A6000666; NID:G2621504; PIDD:AAB84947.1; PTD:G262150  
A/Experimental source: strain Delta H  
C/Genetics:  
A/Gene: MTH441  
A/Start codon: TTG  
C/Superfamily: exonuclease ABC chain C

Query Match 11.2%; Score 81; DB 2; Length 579;  
Best Local Similarity 27.6%; Pred. No. 3.3;  
Matches 43; Conservative 26; Mismatches 53; Indels 34; Gaps 9;

QY 7 RDSALTALDEETL-----MEMES---HRRIVRCICPSRLTPYLQ---AKYLQOLD 53  
DB 262 RDGKITGDPTLNGSAPRTILEAFKQYVALPRVASEILTYPYVDGVALEWLSLR 321  
QY 54 EEEV-IHSP-----RLTNSAMRAGHLDLKTGKNGALFLESLEKPHNDVTVTLVGL 106  
DB 322 GEEVYHSPGAGRRLLINIAWKXASVILKQKGVRRDALQDKDKLKL--PEIPRRMEGL 379

QY 107 QPDVDFSNFSGESSD-----FDG--LAGTSRNLRL 134  
DB 380 ----DISNIAGSATGSAVAFIDGKPSGSGSYRRYRI 411

## RESULT 2

B96505  
hypothetical protein T7023.25 [imported] - Arabidopsis thaliana

C/Species: Arabidopsis thaliana (mouse-ear cress)  
C/Date: 02-Mar-2001 #sequence\_revision 02-Mar-2001 #text\_change 31-Mar-2001

C/Accession: B96505

R/Theologis, A.; Becker, J.R.; Palm, C.J.; Federpiel, N.A.; Kaul, S.; White, O.; Alonso,  
Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.;  
ansen, N.F.; Hughes, B.; Hutzar, L.  
Nature 408, 816-820, 2000

A/Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.  
C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Lueros, J.S.; Maltli, R.; Marziani,

Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.  
 A:Authors: Salzman, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,  
 Kerr, M.; Wu, D.; Xu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.  
 A>Title: Sequence and analysis of chromosome 1 of the plant *Arabidopsis*.  
 A:Reference number: A86141; MUID:21016719; PMID:11130712  
 A:Accession: E96505  
 A>Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-2114 <STO>  
 A:Cross-references: GB:AE005173; NID:g11120821; PIDN:AA31000.1; GSPDB:GN00141  
 C:Genetics:  
 A:Gene: T7023.25  
 A:Map position: 1

Query Match 11.1%; Score 80; DB 2; Length 2114;  
 Best Local Similarity 26.5%; Pred. No. 21;  
 Matches 40; Conservative 22; Mismatches 69; Indels 20; Gaps 5;  
 Db 4 LCRDSDALNALDE-ETLWEMESHHR-----YRCICPRLTPYLQAKVLCQL 52  
 1206 LSPDSTETVSELSLSPSPITRHKTAISSMQLGILHLASRTY-NAAVLCGL 1264  
 Qy 53 DEEEVHSPRLTNSAMRAGHLLDLKTRGKGAIAFLSESLK-----FENPPVTLVTGL 106  
 Db 1265 FSEEHRLDEELANKAL--SPLEMTNTLESERVALTLVLTWGINRPDILTSLEGN 1322  
 Qy 107 QPVDVDFSNFSGSSDPDGLAGTSRNRLRLVT 137  
 Db 1323 PLDNIYKILSLDSSLSKTSAAICRIFLT 1353

RESULT 3  
 H83493  
 hypothetical protein PA1214 [imported] - *Pseudomonas aeruginosa* (strain PA01)  
 C:Species: *Pseudomonas aeruginosa*  
 C>Date: 15-Sep-2000 #sequence\_revision 15-Sep-2000 #text\_change 31-Dec-2000  
 C:Accession: H83493  
 R:Stover, C.K.; Pham, X.Q.; Ewlin, A.L.; Mizoguchi, S.D.; Warren, P.; Hickey, M.J.; Br  
 adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kae, A.; Lardig, K.; Lim,  
 .; Lory, S.; Olson, M.V.  
 A>Title: Complete genome sequence of *Pseudomonas aeruginosa* PA01, an opportunistic patho  
 A:Reference number: A82950; MUID:20437337; PMID:10984043  
 A:Accession: H83493  
 A>Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-532 <STO>  
 A:Cross-references: GB:AE004551; GB:AE004091; NID:g9947135; PIDN:AA04603.1; GSPDB:GN001  
 A:Experimental source: strain PA01  
 C:Genetics:  
 A:Gene: PA1214

Query Match 10.6%; Score 76.5; DB 2; Length 532;  
 Best Local Similarity 26.7%; Pred. No. 8.5;  
 Matches 40; Conservative 18; Mismatches 67; Indels 25; Gaps 6;  
 Qy 4 LCRDSDALNALDE-ETLWEMESHHRIRIVC-----ICPSRLTPYLQ 45  
 Db 197 LCRDSDALNALDE-ETLWEMESHHRIRIVC-----ICPSRLTPYLQ 255  
 Qy 46 AKVLCQDEEEVHSPRLTNSAMRAGHLLDLKTRGKGAIAFLSESLKFNHPDVYTLVT 104  
 Db 256 FHLCLGDEESLPHRLQGFELRQAFMLPILRAVNGGATRMSSLLMYG----RLAD 311  
 Qy 105 GLQPDVDFSNFSGSSDPDGLAGTSRNRLRL 134  
 Db 312 GIGGGYHCVTLGEGAD-ELFMGYPRHLEL 340

RESULT 4  
 A89124  
 probable phosphoesterase (EC 3.1.-.-) K07C11.7 [similarity] - *Caenorhabditis elegans*  
 C:Species: *Caenorhabditis elegans*

C>Date: 10-May-2001 #sequence\_revision 10-May-2001 #text\_change 02-Nov-2001  
 C:Accession: A89124  
 R:Anonymous. The *C. elegans* Sequencing Consortium.  
 A>Title: Genome sequence of the nematode *C. elegans*: a platform for investigating biolo  
 Science 282, 2012-2018, 1996  
 A:Reference number: A75000; MUID:99069613; PMID:9851916  
 A>Note: see websites genome.wustl.edu/sec/C.elegans/ and www.sanger.ac.uk/Projects/C.el  
 A>Note: published errata appeared in Science 283, 35, 1999; Science 283, 2103, 1999; am  
 A:Accession: A89124  
 A:Molecule type: DNA  
 A:Residues: 1-271 <STO>  
 A:Cross-references: GB:chr V; PIDN:AA9617.1; PID:g1255827; GSPDB:GN00023; CESP:K07C11  
 C:Comment: This sequence has motifs characteristic of a variety of phosphoesterases.  
 C:Genetics:  
 A:Gene: K07C11.7  
 A:Map position: 5  
 C:Superfamily: *Caenorhabditis elegans* probable phosphoesterase C25E10.12; phosphoesterase  
 C:Keywords: hydrolase  
 F:68-126/Domain: phosphoesterase core homology <PFC>

Query Match 10.4%; Score 75; DB 1; Length 271;  
 Best Local Similarity 26.5%; Pred. No. 5.2;  
 Matches 36; Conservative 19; Mismatches 61; Indels 20; Gaps 6;  
 Qy 14 LDEETLWEMESHHRIRIVC-----PSRLTPYLQAKVLCQDEEEVHSPR 62  
 Db 27 IDPAENELWDSIHRTYQNIVERKALGINSPTDTPYL--KVCISDTHQLNVT 83  
 Qy 63 LTNSAMRAGHLLDLKTRGKGAIAFLSESLKFNHPDVYTLVTGLQPDVDFSN--FSGES 119  
 Db 84 VPDGDVLI-HAGDTNNCKREELIKNEWTRF--PKYKLVAGNHLELFDHDENGER 140  
 Qy 120 SDPDGLAGTSRNRLRL 135  
 Db 141 QDADKGLGTEGNYIL 156

RESULT 5  
 S55671  
 hypothetical protein E10 - equine herpesvirus 2  
 C:Species: equine herpesvirus 2  
 C>Date: 27-Oct-1995 #sequence\_revision 03-Nov-1995 #text\_change 24-Nov-1999  
 C:Accession: S55671  
 R:Teiford, E.A.R.; Watson, M.S.; Aird, H.C.; Perry, J.; Davison, A.J.  
 J. Mol. Biol. 249, 520-528, 1995  
 A>Title: The DNA sequence of equine herpesvirus 2.  
 A:Reference number: S55594; MUID:95302501; PMID:7783207  
 A:Accession: S55671  
 A>Status: preliminary; nucleic acid sequence not shown; translation not shown  
 A:Molecule type: DNA  
 A:Residues: 1-210 <TEL>  
 A:Cross-references: GB:U20824; NID:g695172; PIDN:AA013865.1; PID:g695250  
 A>Note: the nucleotide sequence was submitted to the EMBL Data Library, February 1995  
 C:Superfamily: equine herpesvirus 2 hypothetical protein E10

Query Match 10.3%; Score 74; DB 2; Length 210;  
 Best Local Similarity 28.0%; Pred. No. 4.7;  
 Matches 30; Conservative 14; Mismatches 27; Indels 36; Gaps 6;  
 Qy 2 GELCRDSDALNALDEETLWEM-----MESHRRIRIVRCICPSRLTPYLQ 46  
 Db 10 GDPC-----VTLREEDIMVERCLRELRVLLVSHKSHK-----LDHLRAK 52  
 Qy 47 KVLQCLDEEEVHSPRLTNSAMRAGHLLDLKTRGKGAIAFLSESLK 93  
 Db 53 KILSRDAREV--SSRAI-SRSRAGLLVDMCQDHR-GFOCLKSSCK 95

RESULT 6  
 T26673  
 hypothetical protein Y3810A.w - *Caenorhabditis elegans*  
 C:Species: *Caenorhabditis elegans*  
 C>Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 20-Jun-2000

C/Accession: T26673  
R/Mills, J.  
Submitted to the EMBL Data Library, September 1999  
A/Reference number: Z20252  
A/Accession: T26673  
A/Status: preliminary; translated from GB/EMBL/DBJ  
A/Molecule type: DNA  
A/Residues: 1-315 <WIL>  
A/Cross-references: EMBL:AL110484; PIDN: CAB54414.1; CESP: Y38E10A.W  
A/Experimental source: clone Y38E10A  
C/Genetics:  
A/Gene: CESP: Y38E10A.W  
A/Intons: 112/1; 172/2  
C/Superfamily: Caenorhabditis elegans hypothetical protein Y38E10A.W

Query Match 10.2%; Score 73.5; DB 2; Length 315;  
Best Local Similarity 27.0%; Pred. No. 8.9;  
Matches 31; Conservative 15; Mismatches 42; Indels 27; Gaps 5;

23 MESHRRIVACIPSRLLTPYLROAKVLCQDEEVLHSPRLTNSAMRA-----GHLD 75  
Db 18 LAENRPFQALADSPFTYVTKLKRQAE-----SPRLSALSLAISIVSHLPA 70

QY 76 LKTRGNKAIA---PLESLKFNPDV-----YLVLTGLQPDVDFSNFSG 117  
Db 71 VLKQKNGKGLTAEKLIQSTKDKHKNVKAQRMVSPYLTLYKKNPFVFPADLQ 125

RESULT 7  
T14260  
period protein Per3 - mouse  
C/Species: Mus musculus (house mouse)  
C/Date: 20-Sep-1999 #sequence\_revision 20-Sep-1999 #text\_change 20-Sep-1999  
C/Accession: T14260  
R/Zylka, M.J.; Shearman, L.P.; Weaver, D.R.; Repert, S.M.  
Neuron 20, 1103-1110, 1998  
A/Title: Three period homologs in mammals: differential light responses in the suprachiasmatic nucleus  
A/Reference number: Z17943; MID: 98318231; PMID: 9655499  
A/Accession: T14260  
A/Status: preliminary; translated from GB/EMBL/DBJ  
A/Molecule type: mRNA  
A/Residues: 1-1113 <ZYL>  
A/Cross-references: EMBL:AF050182; NID: G3136149; PID: G3136150; PIDN: AAC40147.1  
C/Genetics:  
A/Gene: Per3  
C/Keywords: circadian rhythm

Query Match 10.2%; Score 73.5; DB 2; Length 1113;  
Best Local Similarity 25.2%; Pred. No. 44;  
Matches 26; Conservative 19; Mismatches 33; Indels 25; Gaps 4;

6 RPSALVALDELLWEMSHRIRVRCICPSRLTYLNGAKTLQ-----LDEEVL 58  
Db 1003 QRBDAIPGALESIMWYIR---RTPECVIMTYQVPERGSEVLKQDLEKLSMEQOQPL 1058

QY 59 HSPRL-----TNSAMRAGHLIDL--KTRGNKAIA 87  
Db 1059 FSRAGREELAKVSMHSHTAPQEGHLSVCACEDSGSVGDTA 1101

RESULT 8  
H97280  
hydroxyethylthiazole kinase TH1M/Th1K (FS1) CAC3095 [imported] - Clostridium acetobutylicum  
C/Species: Clostridium acetobutylicum  
C/Date: 14-Sep-2001 #sequence\_revision 14-Sep-2001 #text\_change 30-Sep-2001  
C/Accession: H97280  
R/Nolling, J.; Bretton, G.; Omeletchenko, M.V.; Markarova, K.S.; Zeng, Q.; Gibson, R.; Lee, J.; Daly, M.J.; Bennett, G.N.; Koonin, E.V.; Smith, D.R.  
J. Bacteriol. 183, 4823-4838, 2001  
A/Title: Genome Sequence and Comparative Analysis of the Solvent-Producing Bacterium Clostridium acetobutylicum  
A/Reference number: A96900; MID: 21359325; PMID: 21359325  
A/Accession: H97280  
A/Status: preliminary

A/Molecule type: DNA  
A/Residues: 1-265 <KUR>  
A/Cross-references: GB:AE001437; PIDN: AAK81035.1; PID: G15026160; GSPDB: GN00168  
A/Experimental source: Clostridium acetobutylicum ATCC824  
C/Genetics:  
A/Gene: CAC3095  
C/Superfamily: phosphomethylpyrimidine phosphate kinase

Query Match 10.1%; Score 73; DB 2; Length 265;  
Best Local Similarity 23.8%; Pred. No. 8.1;  
Matches 20; Conservative 23; Mismatches 29; Indels 12; Gaps 4;

29 RIVRCICP--SPRLPYLROAKVLCQ---DDEEVLHSPRLTNS-----AMRAGHLID-L 76  
Db 125 KLITLLEPGITITNTEAEAVLCGFKKSEODVRAAKKISEKINVGILTKGHLVDNA 184

QY 77 LKTRGNKAIAPLESLKFNPDV 100  
Db 185 TDLFYKNGVTEWFKLEKIDNPNT 208

RESULT 9  
T21266  
hypothetical protein F22E10.1 - Caenorhabditis elegans  
C/Species: Caenorhabditis elegans  
C/Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 18-Feb-2000  
C/Accession: T21266  
R/Gardner, A.  
Submitted to the EMBL Data Library, November 1995  
A/Reference number: Z19398  
A/Accession: T21266  
A/Status: preliminary; translated from GB/EMBL/DBJ  
A/Molecule type: DNA  
A/Residues: 1-1318 <WIL>  
A/Cross-references: EMBL:Z67882; PIDN: CAA31799.1; GSPDB: GN00028; CESP: F22E10.1  
A/Experimental source: clone F22E10  
C/Genetics:  
A/Gene: CESP: F22E10.1  
A/Map position: X  
A/Intons: 26/3; 81/3; 114/2; 156/3; 201/2; 258/3; 300/2; 396/3; 676/2; 818/3; 939/3; 10  
C/Superfamily: unassigned ATP-binding cassette proteins; ATP-binding cassette homology

Query Match 10.1%; Score 72.5; DB 2; Length 1318;  
Best Local Similarity 26.9%; Pred. No. 69;  
Matches 25; Conservative 13; Mismatches 32; Indels 23; Gaps 4;

57 VLHS---PRLTNSAMRAGHLIDLKTRGNKAIA-----AFLESLKFNPDVYT 101  
Db 1032 VHSHPNPFEPVAKTAAGMLFKLIYRSKTDGWEENNTIRGNVLPESVYFSPQ--- 1088

QY 102 LVTGLQPDVDFSNFSGSSDPDGLAGTSRN 131  
Db 1089 --RPMQPVMTDLHFAHSGQVVALVGPSTGKS 1119

RESULT 10  
S59644  
sister chromatid cohesion molecule M184p - fission yeast (Schizosaccharomyces pombe)  
C/Species: Schizosaccharomyces pombe  
C/Date: 14-Jan-1996 #sequence\_revision 19-Apr-1996 #text\_change 18-Feb-2000  
C/Accession: T38603; T43392; S59644  
R/Devlin, K.; Churcher, C.M.; Barrell, B.G.; Randal, M.A.; Walsh, S.V.  
Submitted to the EMBL Data Library, July 1995  
A/Reference number: Z21731  
A/Accession: T38603  
A/Status: preliminary; translated from GB/EMBL/DBJ  
A/Molecule type: DNA  
A/Residues: 1-1583 <DE2>  
A/Cross-references: EMBL:Z50113; NID: G914878; PIDN: CAA30463.1; PID: G914883; GSPDB: GN0006  
A/Experimental source: strain 972h-1; cosmid c312  
R/Furuya, K.; Takahashi, K.; Yanagida, M.  
submitted to the EMBL Data Library, August 1998  
A/Description: Faithful anaphase is ensured by M184, a sister chromatid cohesion molecule

A/Reference number: Z22478  
 A/Accession: T43392  
 A/Status: preliminary; translated from GB/EMBL/DBJ  
 A/Molecule type: DNA  
 A/Residues: 1-1583 <FUR>  
 A/Cross-references: EMBL:AB016866; PDB:1A9489.1  
 C/Genetics:  
 A/Gene: mls4, SPAC31A2.05C  
 A/Map position: 1  
 A/Introns: 33/1, 98/2, 543/3, 699/3, 1294/2, 1339/3, 1558/3

Query Match 10.1%; Score 72.5; DB 2; Length 1583;  
 Best Local Similarity 29.0%; Pred. No. 87;  
 Matches 31; Conservative 14; Mismatches 39; Indels 23; Gaps 6;

11 LTALDEETLWEMSHRRIRVRCIP--SRLTPYRQAKVL--QQLDEEYLAHSRLTNS 66  
 Db 1091 LTXAGTALME-----IVPCLCSLFTRLNDYERLKKIVSCSKSIEARHS--ENN 1139

67 AMRAGHLIDLKTRGNKGA1-----AFLESLEKFNPD--DYVTLVTVG 105  
 Db 1140 FQKXWRLDILGFRXGDLNRIKNDWKHSIDFISPCDDAYVILG 1186

RESULT 11  
 137468  
 house keeping gene 33 - human  
 C/Species: Homo sapiens (man)  
 C/Date: 29-May-1998 #sequence\_revision 29-May-1998 #text\_change 18-Aug-2000  
 C/Accession: 137468  
 R/Braun, A.; Kammerer, S.; Weissenhorn, W.; Weiss, E.H.; Cleve, H.  
 Gene 146, 291-295, 1994  
 A/Title: Sequence of a putative human housekeeping gene (HK33) localized on chromosome 1  
 A/Reference number: 137468; M01D:94357452; PMID:8076834  
 A/Accession: 137468  
 A/Status: preliminary;  
 A/Molecule type: mRNA  
 A/Cross-references: EMBL:X75535; NID:g1644300; PDB:1CA53225.1; PDB:9551250  
 C/Genetics:  
 A/Gene: hk 33  
 C/Superfamily: Caenorhabditis elegans F54F2.8 protein

Query Match 10.0%; Score 72; DB 2; Length 299;  
 Best Local Similarity 26.0%; Pred. No. 12;  
 Matches 32; Conservative 20; Mismatches 51; Indels 20; Gaps 6;

1 MBEICRDSALTADE--ETLWEMSHRRIRVRCIPSRITFPYRQAKVLCQLDEEYVL 58  
 Db 179 MGNLSKDVLYPSLKEITKTPSWLQSHRESLP---PQCFKTYEQHSVMCKICEQEA 234

59 HSPRLTNSAMRAGH--LTDLKTRGNKGAIAFLSKFHPNDVYTVGLQPDVDFSN 114  
 Db 235 EPTDSEITQKARFENVLDLMQ-----LQDLG-HPPKELAGMPPGLNPLDLALN 284

115 FSG 117  
 Db 285 LSG 287

RESULT 12  
 B36186  
 I factor protein 2 - fruit fly (Drosophila teissleri) transposon I  
 C/Species: Drosophila teissleri  
 C/Date: 14-Dec-1990 #sequence\_revision 14-Dec-1990 #text\_change 31-Mar-2000  
 C/Accession: B36186  
 R/Abad, P.; Vaury, C.; Pelisson, A.; Chaboisier, M.C.; Buseau, I.; Bucheton, A.  
 Proc. Natl. Acad. Sci. U.S.A. 86, 8887-8891, 1989  
 A/Title: A long interspersed repetitive element--the I factor of Drosophila teissleri--  
 A/Reference number: A36186; M01D:90046894; PMID:2554335  
 A/Accession: B36186  
 A/Status: preliminary  
 A/Molecule type: DNA

A/Residues: 1-1234 <ABA>  
 A/Cross-references: GB:W28878; NID:g157750; PDB:950316  
 C/Genetics:  
 A/Gene: FlyBase:Dmel/I-element  
 A/Cross-references: FlyBase:Fgn0013017

Query Match 10.0%; Score 72; DB 2; Length 1234;  
 Best Local Similarity 27.1%; Pred. No. 71;  
 Matches 26; Conservative 20; Mismatches 34; Indels 16; Gaps 4;

59 HSPRLTNSAMRAGHIDLKTR--RGKNG-----AIAFLSKL-----FHPNDVYTVGLQ 107  
 Db 979 YSSVLTSSETALIEALLKTRGKGFVDFSDAISIDSKNPNNSFYPRIRSLITQLA 1038

108 PDVDFSNFSGSSDF-----DGLAGTSRLRLVTP 138  
 Db 1039 FKIKIMTPEHSGITGNELADQAKLASNPILVTP 1074

RESULT 13  
 T45569  
 receptor protein kinase-like protein - Arabidopsis thaliana  
 N/Alternate names: protein P11C1.70  
 C/Species: Arabidopsis thaliana (mouse-ear cress)  
 C/Date: 04-Feb-2000 #sequence\_revision 04-Feb-2000 #text\_change 04-Mar-2000  
 C/Accession: T45569  
 R/Bargues, M.; Collado, M.C.; Navarro, P.; Terol, J.; Perez-Alonso, M.; Mewes, H.W.; May  
 submitted to the Protein Sequence Database, December 1999  
 A/Reference number: Z23007  
 A/Accession: T45569  
 A/Status: preliminary  
 A/Molecule type: DNA  
 A/Residues: 1-660 <BAR>  
 A/Cross-references: EMBL:A1132976  
 A/Experimental source: cultivar Columbia; BAC clone P11C1  
 C/Genetics:  
 A/Map position: 3  
 A/Introns: 472/1  
 A/Note: P11C1.70  
 C/Superfamily: unassigned Ser/Thr or Tyr-specific protein kinases; protein kinase homol.

Query Match 9.9%; Score 71.5; DB 2; Length 660;  
 Best Local Similarity 27.0%; Pred. No. 36;  
 Matches 38; Conservative 21; Mismatches 41; Indels 41; Gaps 8;

20 WEMESHRRIRVRCIP-----SRLTPYRQAKVLCQLDEEYLAHSRLTNSAMRAGHL 74  
 Db 66 WRGVDGQDVRVRLIDGVGLRGSFSP-----ETLSRLDQLRVL---SLNNST-SGSIP 116

75 DL-----LK--TRGNKALFLSKFHPNDVYTVGLQPDVDFSNFSGE----- 118  
 Db 117 DLSPLVNLKTLTKSKNFGSLSS-----SILSLRTELIDSPNFPNGEIPSGINALS 170

119 -----SSDPDLGAGTSRLRL 134  
 Db 171 RLSSILNLEFNRNLNGTLPPLNL 191

RESULT 14  
 A90269  
 conserved hypothetical protein (imported) - Sulfolobus solfataricus  
 C/Species: Sulfolobus solfataricus  
 C/Date: 24-May-2001 #sequence\_revision 24-May-2001 #text\_change 24-May-2001  
 C/Accession: A90269  
 R/She, Q.; Singh, R.K.; Confalonieri, F.; Zivanovic, Y.; Allard, G.; Aways, M.U.; Chan  
 Jang, I.; Jeffries, A.C.; Kozera, C.J.; Medina, N.; Peng, X.; Thi Ngoc, H.P.; Redder,  
 arett, R.A.; Ragan, M.A.; Semsen, C.W.; Van der Oost, J.  
 Submitted to GenBank, April 2001  
 A/Description: Sulfolobus solfataricus complete genome.  
 A/Reference number: A90269  
 A/Accession: A90269  
 A/Status: preliminary  
 A/Molecule type: DNA





GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 18, 2004, 03:29:02 ; Search time 24 Seconds

(without alignments)  
272.363 Million cell updates/sec

Title: US-10-032-159A-16

Perfect score: 720  
Sequence: 1 M6ELCRDSALTRALDSEILM.....SDPDGLAGTSNRLRLVTPX 139

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 127863 seqs, 47026705 residues

Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database: SwissProt\_41.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	617.5	85.8	1004	1 CARE_HUMAN	Q9bx16 homo sapien
2	516.5	71.7	999	1 CARE_MOUSE	Q99xf0 homo musc
3	270.5	37.6	1147	1 CARE_HUMAN	Q9bx17 homo sapien
4	252	35.0	1021	1 CARE_MOUSE	P58660 mus muscul
5	246	34.2	1032	1 CARE_HUMAN	Q9bwt7 homo sapien
6	223.5	31.0	536	1 CARE_RAT	Q9ep70 rattus norv
7	215.5	29.9	536	1 CARE_HUMAN	Q9h257 homo sapien
8	83	11.5	233	1 BCL2_HUMAN	Q95999 h b cell ly
9	81	11.2	233	1 BCL2_RAT	Q9qy95 rattus norv
10	81	11.2	579	1 UVRC_METTH	Q26541 methanobact
11	80	11.1	965	1 AMPN_RABIT	P15541 oryctolagus
12	79	11.0	519	1 KCG4_HUMAN	Q8tdn1 homo sapien
13	78	10.8	233	1 BCL2_MOUSE	Q920h7 m b cell ly
14	75.5	10.5	685	1 STM1_HUMAN	Q13586 homo sapien
15	75	10.4	290	1 YMSO_CAEEL	Q21268 caenorhabd
16	75	10.4	467	1 TM11_MOUSE	Q979q2 mus muscul
17	73.5	10.2	1113	1 PER3_MOUSE	Q092f1 mus muscul
18	72.5	10.1	1584	1 MISA_SCHPO	Q092f5 schizosacch
19	72	10.0	299	1 PXF_HUMAN	P40855 homo sapien
20	71	9.9	379	1 Y452_METAC	Q8t111 methanobact
21	70.5	9.8	356	1 GBA2_CAEEL	P22454 caenorhabd
22	70.5	9.8	513	1 PDI_MAIZE	P52588 zeae mays
23	70.5	9.8	2376	1 TAO3_YEAST	P40468 saccharomyc
24	69.5	9.7	1452	1 PTPM_MOUSE	P28828 mus muscul
25	69	9.6	965	1 AMPN_MOUSE	P97449 mus muscul
26	69	9.6	966	1 AMPN_HUMAN	P15144 homo sapien
27	68.5	9.5	532	1 RTCR_ECOLI	P38035 escherichia
28	68.5	9.5	966	1 AMPN_FELICA	P79171 felis silve
29	68	9.4	468	1 TM11_HUMAN	Q96444 homo sapien
30	68	9.4	603	1 PLK1_HUMAN	P53350 homo sapien
31	67.5	9.4	822	1 SYV_HELPJ	Q92k61 helicobacte
32	67.5	9.4	862	1 CAD1_HUMAN	P12830 homo sapien
33	67.5	9.4	1420	1 SRB9_YEAST	P38931 saccharomyc

34	67	9.3	231	1 PYRF_OCEIH	Q9er36 oceanobacil
35	67	9.3	728	1 YV89_YEAST	P47156 saccharomyc
36	67	9.3	867	1 SYA_AQUAE	O67323 aquifex aeo
37	67	9.3	902	1 GCP2_HUMAN	Q9b8j2 homo sapien
38	66.5	9.2	841	1 PSP1_YEAST	P50896 saccharomyc
39	66.5	9.2	2245	1 MYSJ_DICDI	P54697 dictyosteli
40	66	9.2	326	1 PDL1_MOUSE	O70400 mus muscul
41	66	9.2	524	1 HY11_BRAJA	P59385 bradyrhizob
42	66	9.2	578	1 V022_FOWPV	Q915h9 fowlpox vir
43	66	9.2	853	1 NUG2_RHIME	P56914 rhizobium m
44	66	9.2	871	1 SYA_AQUAE	Q9xdm3 aquifex pyr
45	66	9.2	903	1 GCP2_MOUSE	Q92198 mus muscul

## ALIGNMENTS

RESULT 1  
ID CARE\_HUMAN STANDARD; PRT; 1004 AA.  
AC Q9BX16; Q9BX16; 1004 AA.  
DT 28-FEB-2003 (Rel. 41, Last sequence update)  
DT 28-FEB-2003 (Rel. 41, Last sequence update)  
DT 15-SEP-2003 (Rel. 42, Last annotation update)  
DE Caspase recruitment domain protein 14 (CARD-containing MAGUK protein  
2) (Carna 2).  
GN CARD14 OR CARMA2.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=21192234; PubMed=11278692;  
RA Berlin J., Wang L., Guo Y., Jacobson M.D., Poyet J.-L.,  
Srinivasula S.M., Merriam S., Disfeyano P.S., Alnemri E.S.,  
"CARD11 and CARD14 are novel caspase recruitment domain  
(CARD)/membrane-associated guanylate kinase (MAGUK) family members  
that interact with Bcl10 and activate NF-kappaB".  
RT J. Biol. Chem. 276:11877-11882(2001).  
RL [2]  
RN SEQUENCE FROM N.A.  
RX MEDLINE=21255663; PubMed=11356195;  
RA Gaide O., Martinson F., Michau O., Bonnet D., Thome M., Tschopp J.,  
FEB8 Lett. 505:198-198(2001).  
RL [4]  
RN SEQUENCE OF 1-740 FROM N.A.  
RX MEDLINE=22388257; PubMed=12477932;  
RA Strussberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
Klausner R.D., Collins F.S., Wagner L., Shennan C.M., Schuler G.D.,  
Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,  
Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,  
Birnbaum M.J., Ueda T.B., Toshimaru S., Carrinchi P., Prange C.,  
Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,  
Bosak S.A., McGowan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
Richards S., Morley K.C., Hale S., Garcia A.M., Gay L.J., Hilyk S.W.,  
Villalón D.K., Wuzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
Fahy J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,  
Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,  
Butterfield V.S.N., Krzywinski M.I., Skalska U., Smalhus D.E.,  
Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.,  
"Generation and initial analysis of more than 15,000 full-length

```

RT      human and mouse cDNA sequences."
CC      Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
CC      - FUNCTION: Activates NF-kappaB via Bcl10 and IKK. Stimulates the
CC      phosphorylation of Bcl10.
CC      - SUBUNIT: CARD14 and Bcl10 bind to each other by CARD-CARD
CC      interaction.
CC      - SUBCELLULAR LOCATION: Cytoplasmic.
CC      - TISSUE SPECIFICITY: Expressed in placenta. Also detected in Hela
CC      S3 cells, but not in the other cancer cell lines tested.
CC      - SIMILARITY: Contains 1 CARD domain.
CC      - SIMILARITY: Contains 1 PDZ/DHR domain.
CC      - SIMILARITY: Contains 1 guanylate kinase-like domain.
CC      - CAUTION: Supposed to contain a SH3 domain which is not detected by
CC      PROSITE, Pfam or SMART.
CC      -----
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.isb-sib.ch/announce
CC      or send an email to license@isb-sib.ch).
CC      -----
DR      EMBL; AF322642; AAG53403.1; -.
DR      EMBL; AY032927; AAK54453.1; -.
DR      EMBL; BC018142; AAH18142.1; -.
DR      EMBL; BC001326; AAH01326.1; ALT_INIT.
DR      Genew; HGNC:16446; CARD14.
DR      MIM; 607211; -.
DR      GO; GO:0005886; C:plasma membrane; NAS.
DR      GO; GO:0005515; F:protein binding activity; NAS.
DR      GO; GO:0007250; P:activation of NF-kappaB-inducing kinase; NAS.
DR      Interpro; IPR001315; CARD.
DR      Interpro; IPR000619; Guanylate_kin.
DR      SMART; SMO0072; GUKC; 1.
DR      SMART; SMO00228; PDZ; 1.
DR      PROSITE; PS50209; CARD; 1.
DR      PROSITE; PS500856; GUANYLATE_KINASE_1; FALSE_NEG.
DR      PROSITE; PS50052; GUANYLATE_KINASE_2; 1.
DR      PROSITE; PSS0106; PDZ; 1.
DR      Colled coil.
KT      FT DOMAIN 15 107 CARD.
FT DOMAIN 128 409 COLLED COIL (POTENTIAL) .
FT DOMAIN 568 658 PDZ.
FT DOMAIN 858 990 GUANYLATE KINASE.
FT CONFLICT 619 671 DYAESEPLFKFVLDQDTLEAVGLLRVDGFCCLSVKYATD
FT CYRLLDLQLEAK --> SRAPPLISPGILMTGTVAAGSYQAD
FT PSPPRCRGSTIGMNASALSMADVRSNHL (IN REF. 2;
FT AAH01326) .
SQ SEQUENCE 1004 AA; 113299 MW; 7469B8B56BE06073 CRC64;
Query Match 85.8%; Score 617.5; DB 1; Length 1004;
Best Local Similarity 90.4%; Pred. No. 1.1e-53;
Matches 123; Conservative 4; Mismatches 6; Indels 3; Gaps 2
QY 1 MGELCRDSALTALDEETLMEMESHRRIRVCIPSLTPPYLRQAAYLCOLDEEEVLHS 60
Db 1 MGELCRDSALTALDEETLMEMESHRRIRVCIPSLTPPYLRQAAYLCOLDEEEVLHS 60
QY 61 PRUTNSAMRAQHILLDLTKRGNGAIALLESIKHPNDPVYTLVTGLQPDVDFENSG--E 118
Db 61 PRUTNSAMRAQHILLDLTKRGNGAIAFLLESIKHPNDPVYTLVTGLQPDVDFENSG-LME 120
QY 119 SSDP-DGLAGTSRNLR 133
Db 121 TSRLTECLAGALSGLQ 136

```

DT 28-FEB-2003 (Rel. 41, Created)  
DT 28-FEB-2003 (Rel. 41, Last sequence update)  
DT 15-SEP-2003 (Rel. 42, Last annotation update)  
DE Gsnpae recruitment domain protein 14 (Bcl10-interacting MAGUK protein  
DE (Bimp2).  
DE  
DE CARD14 OR BIMP2.  
GN Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Euteheria; Rodentia; Scturognathi; Muridae; Murinae; Mus.  
OX NCBI\_TaxID=10090;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=21391892; Pubmed=11387339;  
RA Mallister-Lucas L.M., Inohara N., Lucas P.C., Rulland J., Benito A.,  
RA Li Q., Chen S., Chen F.F., Yamaoka S., Verma I.M., Mak T.W.,  
RA Nunez G.,  
RT "Bimp2, a MAGUK family member linking protein kinase C activation to  
RT Bcl10-mediated NF-kappa B induction."  
RT J. Biol. Chem. 276:30589-30597(2001).  
RN [2]  
RP SEQUENCE OF 82-743 FROM N.A.  
RC TISSUE=Breast;  
RX MEDLINE=22388257; Pubmed=12477932;  
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,  
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Shih N.K.,  
RA Hopkins R.F., Jordan H., Moore T., Wax S.I., Wang J., Heien F.,  
RA Datchenko L., Marusina K., Farmer A.A., Rubin G.M., Heien F.,  
RA Stepien M., Soares M.B., Donald M.F., Casavant T.L., Scheetz T.E.,  
RA Brownstein M.J., Udell T.B., Toshiyuki S., Carninci P., Prange C.,  
RA Raha S.S., Loggiano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,  
RA Bosak S.A., McEwan P.U., McKernan K.J., Malek J.A., Guarracine P.H.,  
RA Richards S., Woley K.C., Hale S., Garcia A.M., Gay L.U., Hulik S.W.,  
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
RA Fahey J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,  
RA Whitting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
RA Blakeley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,  
RA Buttefield Y.S.N., Krzywinski M.I., Skalska U., Smalins D.E.,  
RA Scheraga A., Schein J.E., Jones S.U.M., Marra M.A.,  
RT "Generation and initial analysis of more than 15,000 full-length  
RT human and mouse cDNA sequences".  
RL Proc. Natl. Acad. Sci. U.S.A. 99:16699-16903(2002).  
CC -1- FUNCTION: Activates NF-kappaB via Bcl10 and IKK.  
CC -1- SUBUNIT: CARD14 and Bcl10 bind to each other by CARD-CARD  
CC interaction (By similarity).  
CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).  
CC -1- SIMILARITY: Contains 1 CARD domain.  
CC -1- SIMILARITY: Contains 1 PDZ/DHR domain.  
CC -1- SIMILARITY: Contains 1 guanylate kinase-like domain.  
CC -1- CAUTION: Supposed to contain a SH3 domain which is not detected by  
CC PROSITE, Pfam or SMART.  
-----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce>)  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
-----  
CC  
DR EMBL, AF363457; AA060137.1; -  
DR EMBL, BC004692; AA04692.1; -  
DR MGD, MGI:2386258; Card14.  
DR InterPro: IPR001315; CARD.  
DR InterPro: IPR000619; Guanylate\_kin.  
DR InterPro: IPR001478; PDZ.  
DR Pfam, PF00595; PDZ; 1.  
DR SMART, SM00228; PDZ; 1.  
DR PROSITE, PS50209; CARD; 1.  
DR PROSITE, PS00852; GUANYLATE\_KINASE\_2; 1.  
DR PROSITE, PS50106; PDZ; 1.

KW Coiled coil. 15 CARD.  
 FT DOMAIN 125 411 COILED COIL (POTENTIAL).  
 FT DOMAIN 572 655 PDZ.  
 FT DOMAIN 854 986 GUANYLATE KINASE.  
 FT CONFLICT 736 743 QAOQOLIA -> HLEBDHES (IN REF. 2).  
 SQ SEQUENCE 999 AA; 113496 MW; D18350DA12430255 CRC64;  
 Query Match 71.7%; Score 516.5; DB 1; Length 999;  
 Best Local Similarity 77.2%; Pred. No. 1,3e-43;  
 Matches 105; Conservative 9; Mismatches 19; Indels 3; Gaps 2;  
 QY 1 MGLICRDSALZALDEBTLMEMESHRIRIVRCICPSRLTPYLRQAKVLCQDEEVLHS 60  
 DB 1 MALICMDSTLTLDDEMLMDLHSHRCRIVOSICSRITPYLRQAKVLCQDEEVLHS 60  
 QY 61 PLITNSAMRAGHLIDLKTRGNKGAIAFESLKFENPDVYTLVTLGQPDVDSNFGSG--E 118  
 DB 61 SREFTNSAMRAGHLIDLKTRGNKGAIAFESLKFENPDVYTLVTLGQPDVDSNFGSG 120  
 QY 119 SSDF-DGLAGTSKMLR 133  
 DB 121 TSKLTRECLAGALISLQ 136  
 RESULT 3  
 CARB\_HUMAN  
 ID CARB\_HUMAN STANDARD; PRT; 1147 AA.  
 AC Q9BXL7;  
 DT 28-FEB-2003 (Rel. 41, Created)  
 DT 28-FEB-2003 (Rel. 41, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE Caspase recruitment domain protein 11 (CARD-containing MAGUK protein 3) (Carna 1).  
 GN CARD11 OR CARMA1.  
 OS Homo sapiens (human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 OC NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=21192234; PubMed=11278692;  
 RA Berlin U., Wang L., Guo Y., Jacobson M.D., Poyet J.-L.,  
 RA Srinivasula S.M., Merriam S., Distefano P.S., Alnemri E.S.;  
 RT "CARD11 and CARD14 are novel caspase recruitment domain  
 (CARD)/membrane-associated guanylate kinase (MAGUK) family members  
 that interact with Bcl10 and activate NF-kappaB.";  
 RT J. Biol. Chem. 276:11877-11882(2001).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=21255663; PubMed=11356195;  
 RA Gaide O., Martinon F., Micheau O., Bonnet D., Thome M., Tschopp J.,  
 RA "Carnal, a CARD-containing binding partner of Bcl10, induces Bcl10  
 phosphorylation and NF-kappaB activation.";  
 RT FEBS Lett. 496:121-127(2001).  
 RN [3]  
 RP ERRATUM.  
 RA Gaide O., Martinon F., Micheau O., Bonnet D., Thome M., Tschopp J.,  
 RA FEBS Lett. 505:198-199(2001).  
 CC -1- FUNCTION: Activates NF-kappaB via Bcl10 and IKK. Stimulates the  
 phosphorylation of Bcl10.  
 CC -1- SUBUNIT: CARD11 and Bcl10 bind to each other by CARD-CARD  
 interaction.  
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic.  
 CC -1- TISSUE SPECIFICITY: Detected in adult peripheral blood leukocytes,  
 thymus, spleen and liver. Also found in promyelocytic leukemia HL-  
 60 cells, chronic myelogenous leukemia K562 cells, Burkitt's  
 lymphoma Raji cells and colorectal adenocarcinoma SW620 cells. Not  
 detected in HeLa S3, Molt-4, A549 and G431 cells.  
 CC -1- SIMILARITY: Contains 1 CARD domain.  
 CC -1- SIMILARITY: Contains 1 PDZ/DHR domain.  
 CC -1- SIMILARITY: Contains 1 guanylate kinase-like domain.  
 CC -1- CAUTION: Supposed to contain a SH3 domain which is not detected by

PROSITE, Pfam or SMART.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC EMBL; AF322641; AAC53402.1; -  
 CC Genew; HGNC:16393; CARD11.  
 DR MIM: 607210; -  
 DR GO; GO:0005624; C:membrane fraction; NAS.  
 DR GO; GO:0004384; F:membrane-associated guanylate kinase; NAS.  
 DR GO; GO:0005515; F:protein binding activity; IPI.  
 DR GO; GO:0007250; P:activation of NF-kappaB-inducing kinase; NAS.  
 DR InterPro; IPR001315; CARD.  
 DR InterPro; IPR000619; Guanylate\_kin.  
 DR InterPro; IPR001478; PDZ.  
 DR SMART; SMO0228; PDZ; 1.  
 DR PROSITE; PS50209; CARD; 1.  
 DR PROSITE; PS00856; GUANYLATE\_KINASE\_1; FALSE\_NEG.  
 DR PROSITE; PS50052; GUANYLATE\_KINASE\_2; FALSE\_NEG.  
 DR PROSITE; PS50106; PDZ; FALSE\_NEG.  
 KW Coiled coil.  
 FT DOMAIN 11 103 CARD.  
 FT DOMAIN 123 442 COILED COIL (POTENTIAL).  
 FT DOMAIN 673 748 PDZ.  
 FT DOMAIN 966 1133 GUANYLATE KINASE.  
 FT CONFLICT 808 808 P -> L (IN REF. 2).  
 SQ SEQUENCE 1147 AA; 132641 MW; 91344B015D2B36CC CRC64;  
 Query Match 37.6%; Score 270.5; DB 1; Length 1147;  
 Best Local Similarity 46.6%; Pred. No. 5.1e-19;  
 Matches 54; Conservative 23; Mismatches 33; Indels 1; Gaps 1;  
 QY 15 DEBTLMEMESHRIRIVRCICPSRLTPYLRQAKVLCQDEEVLHSPRLTNSAMRAGHL 74  
 DB 11 EBDALMEVCEGRNMLSYINPAKLTIPYLRQCKYIDGDEDEVLNAPVLPKIRAGRL 70  
 QY 75 DLKTRGNKGAIAFESLKFENPDVYTLVTLGQPDVDSNFGSGSDPDGL 125  
 DB 71 DILTKGQGVVFLSHFYFPHYLYKLVTKSEPTTRFSTIVEEG-HEGL 120  
 RESULT 4  
 CARA\_MOUSE  
 ID CARA\_MOUSE STANDARD; PRT; 1021 AA.  
 AC P58660;  
 DT 28-FEB-2003 (Rel. 41, Created)  
 DT 28-FEB-2003 (Rel. 41, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE Caspase recruitment domain protein 10 (Bcl10-interacting MAGUK protein 1) (Bim1).  
 GN CARD10 OR BIM1.  
 OS Mus musculus (mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OC NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=21391892; PubMed=11387339;  
 RA McAllister-Lucas L.M., Inohara N., Lucas P.C., Ruland J., Benito A.,  
 RA Li Q., Chen S., Chen F.F., Yamaoka S., Verma I.M., Mak T.W.,  
 RA Nunez G.;  
 RT "Bim1, a MAGUK family member linking protein kinase C activation to  
 Bcl10-mediated NF-kappa B induction.";  
 RT J. Biol. Chem. 276:30589-30597(2001).  
 CC -1- FUNCTION: Activates NF-kappaB via Bcl10 and IKK.  
 CC -1- SUBUNIT: CARD10 and Bcl10 bind to each other by CARD-CARD  
 interaction. They both participate in a complex with MLT1, where  
 MULTI binds to Bcl10.

CC -1- TISSUE SPECIFICITY: Highly expressed in kidney, heart followed by  
 CC brain, lung, liver, skeletal muscle and testis.  
 CC -1- SIMILARITY: Contains 1 CARD domain.  
 CC -1- CAUTION: Supposed to contain a SH3, a PDZ and a guanylate kinase-  
 CC like domain. But none of these 3 domains are detected by PROSITE,  
 CC Pfam or SMART.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL: AF363456; AK60136.1; --  
 CC GDB: M61246012; Card10.  
 CC GO: GO:0005737; Cytoplasm; ISS.  
 CC GO: GO:0005515; Protein binding activity; ISS.  
 CC GO: GO:0003015; Protein signaling complex scaffold protein; ISS.  
 CC GO: GO:0007250; Phosphorylation of NF-kappa-inducing kinase; ISS.  
 CC GO: GO:0004651; Protein complex assembly; ISS.  
 CC InterPro: IPR001315; CARD.  
 CC PROSITE: PS50209; CARD; 1.  
 CC Coiled coil.  
 CC KM DOMAIN 23 115 CARD  
 CC FT DOMAIN 138 450 COILED COIL (POTENTIAL).  
 CC FT DOMAIN 558 565 POLY-SER.  
 CC SQ SEQUENCE 1021 AA; 114413 MW; 4811A09BDBF792C CRC64;  
 CC -----  
 CC Query Match 35.0%; Score 252; DB 1; Length 1021;  
 CC Best Local Similarity 51.1%; Pred. No. 3.1e-17;  
 CC Matches 48; Conservative 20; Mismatches 26; Indels 0; Gaps 0;  
 CC -----  
 CC QY 15 DEETLMMESRRHIVRCICPSRLPYLRQAKVLCQDESEVLSPLTNSAPRAGILL 74  
 CC Db 23 EEDALMEREGVRHRLTRALNPAKLTPIRQRCRVLEDEDESEVLSYFPCCRAKRTGLI 82  
 CC QY 75 DLKTRGNKGAIAFLPSIKFNHNDVYTLVTGLOP 108  
 CC Db 83 DILRCRKGKGFAPLEALFEYFPHFTLLTGP 116  
 CC -----  
 CC RESULT 5  
 CC CARD\_HUMAN STANDARD; PRT; 1032 AA.  
 CC AC Q9BWT7; O9UGR5; O9UGR6; O9Y3H0;  
 CC DT 28-FEB-2003 (Rel. 41, Created)  
 CC DT 28-FEB-2003 (Rel. 41, Last sequence update)  
 CC DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 CC DE Caspase recruitment domain protein 10 (CARD-containing MAGUK protein  
 CC 3) (Carda 3).  
 CC GN CARD10 OR CARMA3.  
 CC OS Homo sapiens (Human)  
 CC OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 CC OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
 CC NCBI\_TaxID=9606;  
 CC RN [1]  
 CC RP SEQUENCE FROM N.A.  
 CC RX MEDLINE=21292987; PubMed=11259443;  
 CC RA Wang L., Guo Y., Huang W.-D., Ke X., Poyet J.-L., Manji G.A.,  
 CC Meriam S., Glucksmann M.A., Distefano P.S., Alnemri B.S., Bertin J.,  
 CC "CARD10 is a novel caspase recruitment domain/membrane-associated  
 CC guanylate kinase family member that interacts with Bcl10 and activates  
 CC NF-kappa B.";  
 CC RT J. Biol. Chem. 276:21405-21409(2001).  
 CC RN [2]  
 CC RP SEQUENCE FROM N.A.  
 CC RX MEDLINE=21255663; PubMed=11356195;  
 CC RA Gaido O., Martinon F., Micheau O., Bonnet D., Thome M., Tschopp J.,  
 CC "Cardinal, a CARD-containing binding partner of Bcl10, induces Bcl10  
 CC phosphorylation and NF-kappaB activation.";  
 CC RT FEBS Lett. 496:121-127(2001).  
 CC -----

RN [3]  
 RP ERRATUM.  
 RA Gaido O., Martinon F., Micheau O., Bonnet D., Thome M., Tschopp J.,  
 RL FEBS Lett. 505:198-198(2001).  
 RN [4]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=20057165; PubMed=10591208;  
 RA Dunham I., Hunt A.R., Collins J.E., Bruskewich R., Beare D.M.,  
 RA Clamp M., Smink L.J., Ainscough R., Almeida J.P., Babbage A.K.,  
 RA Baguley C., Bailey J., Barlow K.F., Bates K.N., Beasley O.P.,  
 RA Bird C.P., Blakey S.E., Bridgeman A.M., Buck D., Burgess J.,  
 RA Burrill W.D., Burton J., Carder C., Carter N.P., Chen Y., Clark G.,  
 RA Clegg S.M., Cobley V.E., Cole C.G., Collier R.E., Connor R.,  
 RA Conroy D., Corby N.R., Coville G.J., Cox A.V., Davis J., Dawson B.,  
 RA Dhani P.D., Dockree C., Dodsworth S.J., Durbin R.M., Ellington A.G.,  
 RA Evans K.L., Fey J.M., Fleming K., French L., Garner A.A.,  
 RA Gilbert J.G.R., Goward M.E., Grafham D.V., Griffiths M.N.D., Hall C.,  
 RA Hall R.E., Hall-Tamlyn G., Heathcote R.W., Ho S., Holmes S.,  
 RA Hunt S.E., Jones M.C., Kershaw J., Kimberley A.M., King A.,  
 RA Laird G.K., Langford C.F., Leversha M.A., Lloyd C., Lloyd D.M.,  
 RA McLaren I.D., Maheggh-Mohammadi M., Matthews L.H., Mccann O.T.,  
 RA McElay J., McLaren S., McMurray A.A., Milne S.A., Mortimore B.J.,  
 RA Odell C.N., Pavitt R., Pearce A.V., Pearson D., Phillimore B.J.C.T.,  
 RA Phillips S.H., Plumb R.W., Ramsay H., Ramsey Y., Rogers L., Ross M.T.,  
 RA Scott C.E., Sehra H.K., Skuce C.D., Smalley S., Smith M.L.,  
 RA Soderlund C., Spragon L., Steward C.A., Sulston J.E., Swann R.M.,  
 RA Vaudin M., Wall M., Wallis J.M., Whiteley M.N., Willey D.L.,  
 RA Williams L., Williams S.A., Williamson H., Wilmer T.E., Wilming L.,  
 RA Wright C.L., Hubbard T., Bentley D.R., Beck S., Rogers J., Shimizu N.,  
 RA Minoshima S., Kawasaki K., Sasaki T., Asakawa S., Kudoh J.,  
 RA Shintani A., Shibuya K., Yoshizaki Y., Aoki N., Mitsuyma S.,  
 RA Roe B.A., Chen F., Chu L., Crabtree J., Deschamps S., Do A., Do T.,  
 RA Dorman A., Fang F., Fan Y., Hu P., Hua A., Kenton S., Lai H., Lao H.I.,  
 RA Lewis J., Lewis S., Lin S.-P., Loh P., Mai J., Mayaj E., Nguyen T., Pan H.,  
 RA Pham S., Qi S., Qian Y., Ray L., Ren O., Shaili S., Sloan D., Song L.,  
 RA Wang C., Wang Y., Wang Z., White J., Williamson D., Wu H., Yao Z.,  
 RA Zhan M., Zhang G., Chisoe S., Murray J., Miller N., Mink P.,  
 RA Fulton R., Johnson D., Bemis G., Bentley D., Bradshaw H., Bourne S.,  
 RA Cordes M., Du Z., Fulton L., Goela D., Graves T., Hawkins J.,  
 RA Hinds K., Kemp K., Latreille P., Layman D., Ozersky P., Roloff T.,  
 RA Scheet P., Walker C., Mamsley A., Mohlmann P., Pepin K., Nelson R.,  
 RA Korfi I., Bedell J.A., Hillier L., Wards E., Waterson R., Wilson R.,  
 RA Emanuel B.S., Shaikh T., Kurahashi H., Salita S., Budarf M.L.,  
 RA McDermid H.E., Johnson A., Wong A.C.C., Morrow B.E., Belman L.,  
 RA Kim U.J., Shizuya H., Simon M.I., Dumanski J.P., Peyrard M., Kedra D.,  
 RA Seroussi E., Franssion I., Tapia I., Bruder C.E., O'Brien K.P.,  
 RA Wilkinson P., Boderech A., Hartman K., Hu X., Khan A.S., Lane L.,  
 RA Tliahun Y., Wright H.,  
 RT "The DNA sequence of human chromosome 22.";  
 RL Nature 402:489-495(1999).  
 CC -1- FUNCTION: Activates NF-kappaB via Bcl10 and IKK.  
 CC -1- SUBUNIT: CARD10 and Bcl10 bind to each other by CARD-CARD  
 CC interaction. They both participate in a complex with MALTI, where  
 CC MALTI binds to Bcl10 (by similarity).  
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).  
 CC -1- TISSUE SPECIFICITY: Detected in adult heart, kidney and liver;  
 CC lower levels in intestine, placenta, muscle and lung. Also found  
 CC in fetal lung, liver and kidney.  
 CC -1- SIMILARITY: Contains 1 CARD domain.  
 CC -1- CAUTION: Supposed to contain a SH3, a PDZ and a guanylate kinase-  
 CC like domain. But none of these 3 domains are detected by PROSITE,  
 CC Pfam or SMART.  
 CC -1- CAUTION: Ref. 4 sequence differs from that shown due to various  
 CC gene identification problems.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----

```

DR EMBL; AY028896; AAK26165.1; -.
DR EMBL; AY032928; AAK54454.1; -.
DR EMBL; AL049851; CAB63075.1; ALT_SEQ.
DR EMBL; AL049851; CAB63076.1; ALT_SEQ.
DR EMBL; AL022315; CAB42832.1; ALT_SEQ.
DR Genew; H9NC; 16422; CARD10.
DR MIM; 607209; -.
DR GO; GO:0005737; C:cytoplasm; NAS.
DR GO; GO:0005555; P:protein binding activity; IPT.
DR GO; GO:0030159; P:receptor signaling complex scaffold protein. .; NAS.
DR GO; GO:0007250; P:activation of NF-kappa-inducing kinase; IDA.
DR GO; GO:0006461; P:protein complex assembly; NAS.
DR InterPro; IPR001315; CARD.
DR PROSITE; PS50209; CARD; 1.
KM Coiled coil.
FT DOMAIN 23 115 CARD.
FT DOMAIN 138 456 COILED COIL (POTENTIAL).
FT DOMAIN 567 574 POLY-SER.
FT CONFLICT 289 289 Q -> R (IN REF. 4).
FT CONFLICT 917 917 K -> KQ (IN REF. 4; CAB63075).
FT CONFLICT 932 932 R -> L (IN REF. 4).
SQ SEQUENCE 1032 AA; 115946 MW; 8377319AB82A0949 CRC64;

Query Match 34.2%; Score 246; DB 1; Length 1032;
Best Local Similarity 46.9%; Pred. No. 1.3e-16;
Matches 46; Conservative 21; Mismatches 27; Indels 0; Gaps 0;

QY 15 DEFTLMMESHRRIRVRCICPSRLTPYLRQAKVLCQDEEVLHSPRLTNSMRAHLL 74
DB 23 EDLALMERIBGVKRLARALNPAKLTLYLRQCRVIDQDEEVLSTRFCRVNRTGRLL 82
QY 75 DLKTRGNKGAIAFLSLKFNPDVYTLVTGLOP 108
DB 83 DILRCGRGVYAVFLALFLFYPHFTLLTQGP 116

RESULT 6
CAR9_HUMAN STANDARD; PRT; 536 AA.
AC OSEPYO;
ID CAR9_RAT
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Caspase recruitment domain protein 9 (CARD9).
GN CARD9.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
NCBI_TaxId=10116;
[1]
SEQUENCE FROM N.A.
RC STRAIN=Sprague-Dawley;
RX MEDLINE=20576268; PubMed=11053425;
RA Bertin J., Guo Y., Wang L., Srinivasula S.M., Jacobson M.D.,
RA Poyet J.-L., Meriam S., Du M.-Q., Dyer M.J.S., Robison K.E.,
RA Distefano P.S., Alnemri E.S.;
RA "CARD9 is a novel caspase recruitment domain-containing protein that
interacts with Bcl10/CLAP and activates NF-kappa B."
RT J. Biol. Chem. 275:41082-41086(2000).
RL
-1- FUNCTION: Activates NF-kappaB via Bcl10 (By similarity).
-1- SUBUNIT: Self-associates. CARD9 and Bcl10 bind to each other by
CARD-CARD interaction (By similarity).
-1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
-1- SIMILARITY: Contains 1 CARD domain.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL Outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.ebi.ac.uk/announcements
or send an email to license@ebi.ac.uk).
```

```

DR EMBL; AF311288; AAG28791.1; -.
DR InterPro; IPR001315; CARD.
DR PROSITE; PS50209; CARD; 1.
KM Coiled coil.
FT DOMAIN 6 98 CARD.
FT DOMAIN 117 277 COILED COIL (POTENTIAL).
FT DOMAIN 303 420 COILED COIL (POTENTIAL).
SQ SEQUENCE 536 AA; 62631 MW; 6F33089CB7E6BAC9 CRC64;

Query Match 31.0%; Score 223.5; DB 1; Length 536;
Best Local Similarity 45.4%; Pred. No. 1e-14;
Matches 49; Conservative 18; Mismatches 38; Indels 3; Gaps 1;

QY 15 DEFTLMMESHRRIRVRCICPSRLTPYLRQAKVLCQDEEVLHSPRLTNSMRAHLL 74
DB 6 NDECSWALSFPVKLISVIDSPSRTIPYLRQCKVNPDEEVLDPNIVIRKRGVLL 65
QY 75 DLKTRGNKGAIAFLSLKFNPDVYTLVTGLOPVDPS--NFGES 119
DB 66 DILQTRGHKGVYAVFLSLKFNPDVYTLVTGLOPVDPS--NFGES 113

RESULT 7
CAR9_HUMAN STANDARD; PRT; 536 AA.
AC Q9H257; Q9H854;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 15-FEB-2003 (Rel. 42, Last annotation update)
DE Caspase recruitment domain protein 9 (CARD9).
GN CARD9.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
NCBI_TaxId=9606;
[1]
SEQUENCE FROM N.A.
RX MEDLINE=20576268; PubMed=11053425;
RA Bertin J., Guo Y., Wang L., Srinivasula S.M., Jacobson M.D.,
RA Poyet J.-L., Meriam S., Du M.-Q., Dyer M.J.S., Robison K.E.,
RA Distefano P.S., Alnemri E.S.;
RA "CARD9 is a novel caspase recruitment domain-containing protein that
interacts with Bcl10/CLAP and activates NF-kappa B."
RT J. Biol. Chem. 275:41082-41086(2000).
RL
[2]
SEQUENCE FROM N.A.
RC TISSUE=Retinoblastoma;
RA Iisogai T., Ota T., Hayashi K., Sugiyama T., Otsuki T., Suzuki Y.,
RA Nishikawa T., Nagai K., Sugano S., Shiratori A., Sudo H.,
RA Wagaatsuma M., Hosoi T., Kaku Y., Kodaira H., Kondo H., Sugawara M.,
RA Takahashi M., Chiba Y., Ishida S., Murakawa K., Ono Y., Takiguchi S.,
RA Watanabe S., Kimura K., Murakami K., Ishii S., Kawai Y., Saito K.,
RA Yamamoto K., Wakamatsu A., Nakamura Y., Nagahara K., Masuno Y.,
RA Minomiya K., Iwamoto T.;
RA "NEDD human cDNA sequencing project."
RT Submitted (Aug-2000) to the EMBL/Genbank/DBJ databases.
RL
[3]
SEQUENCE OF 1-492 FROM N.A.
RC TISSUE=Muscle;
RX MEDLINE=22388257; PubMed=12477932;
RA Krausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shemen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Szelepen M., Soares M.B., Bonaldo M.F., Casavant T.L., Schetz T.E.,
RA Brownstein M.U., Uediri T.B., Toshiyuki S., Carrinci P., Prange C.,
RA Rana S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Murty D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahy J., Helton E., Ketterman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting W., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
```

RA Blakeley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
 RA Rodriguez A.C., Grimwood J., Schmitt J., Myers R.W.,  
 RA Butlerfield Y.S.N., Krzywdinski M.I., Skalska U., Smailus D.E.,  
 RA Scheraga A., Schein J.E., Jones S.J.M., Marra M.A.;  
 RT "Generation and initial analysis of more than 15,000 full-length  
 RT human and mouse cDNA sequences";  
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).  
 CC -1- FUNCTION: Activates NF-kappaB via Bcl10.  
 CC -1- SUBUNIT: Self-associates. CARD9 and Bcl10 bind to each other by  
 CC CARD-CARD interaction.  
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic.  
 CC -1- TISSUE SPECIFICITY: Highly expressed in spleen. Also detected in  
 CC liver, placenta, lung, peripheral blood leukocytes and in brain.  
 CC -1- SIMILARITY: Contains 1 CARD domain.  
 CC -1- CAUTION: Ref.2 sequence differs from that shown due to a  
 CC frameshift in position 360.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; AF31287; AAG28790.1; -;  
 DR EMBL; AK024001; BAB14766.1; ALT\_FRAME.  
 DR EMBL; BC008877; AAH08877.1; -;  
 DR Genew; H9NC;16391; CARD9.  
 DR MIM; 607212; -;  
 DR InterPro; IPR001315; CARD.  
 DR PROSITE; PS50209; CARD; 1.  
 KM Coiled coil.  
 FT DOMAIN 6 98  
 FT DOMAIN 117 277  
 FT DOMAIN 332 419  
 FT CONFLICT 12 12 N -> S (IN REF. 3).  
 FT CONFLICT 482 492 LSGEPPERK -> PARGPGGAVC (IN REF. 3).  
 SQ SEQUENCE 536 AA; 62267 MW; 6EB1835315B83DE5 CR664;  
 QY Query Match 29.9%; Score 215.5; DB 1; Length 536;  
 Best Local Similarity 44.4%; Pred. No. 6.4e-14;  
 Matches 48; Conservative 17; Mismatches 40; Indels 3; Gaps 1;  
 QY 15 DEETLEMEHSHRHRIVRCIPRLTPYLRQAKVLCQDEEVLHSPRITSAMRAGHLL 74  
 DB 6 NDECEWVVEGFRVLTSLVSDPRLTPYLRQCKVLPDDEEQLSPNVLIRKRVGL 65  
 QY 75 DLKKTGKXGATAFESLKFHNDVTLVTGLOPDVDS--NFGDS 119  
 DB 66 DILQRTGHKGYAFLESLELYYPOLKYKVGKPEAFVFSMIIDAGS 113  
 RESULT 8  
 BCL10\_HUMAN STANDARD; PRT; 233 AA.  
 ID BCL10\_HUMAN STANDARD; PRT; 233 AA.  
 AC 095999;  
 DT 28-FEB-2003 (Rel. 41, Created)  
 DT 28-FEB-2003 (Rel. 41, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE B cell lymphoma/leukemia 10 (B-cell CLL/lymphoma 10) (Bcl-10) (CED-  
 DE 3/ICH-1 prodomain homologous E10-like regulator) (CIPEB) (CARD-  
 DE containing molecule enhancing NF-kappaB) (cellular homolog of VCAMEN)  
 DE (VCAMEN) (Mammalian CARD-containing adapter molecule E10) (ME10)  
 DE (cellular-E10) (C-E10) (CARD-like apoptotic protein) (hCLAP).  
 GN BCL10 OR CIPEB OR CLAP.  
 OS Homo sapiens (human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.  
 OX NCBI\_TaxId=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A., VARIANT FOLLICULAR LYMPHOMA GLU-210 DEL, VARIANT  
 RP MESOTHELIOMA ILE-52, AND VARIANTS GERM CELL TUMOR GLY-58 AND PHE-219.

RC TISSUE=Lymphoma;  
 RX MEDLINE=99142601; PubMed=9989495;  
 RA Willis T.G., Jadayel D.M., Du M.-Q., Peng H., Perry A.R.,  
 RA Abdul-Rauf M., Price H., Karran L., Matekodomu O., Wlodarska I.,  
 RA Paul L., Crook T., Hamoudi R., Isaacson P., Dyer M.J.S.;  
 RT "Bcl10 is involved in t(1;14)(p22;q32) of MALT B cell lymphoma and  
 RT mutated in multiple tumor types.";  
 RL Cell 96:35-45(1999).  
 RN [2]  
 RP SEQUENCE FROM N.A., AND MUTAGENESIS OF LEU-41 AND GLY-78.  
 RX MEDLINE=99214545; PubMed=10187770;  
 RA Koseki T., Inohara N., Chen S., Carrio R., Merino J., Hottiger M.O.,  
 RA Nabel G.J., Nunez G.;  
 RT "CIPEB, a novel NF-kappaB-activating protein containing a caspase  
 RT recruitment domain with homology to Herpesvirus-2 protein E10.";  
 RL J. Biol. Chem. 274:9955-9961(1999).  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=99214546; PubMed=10187771;  
 RA Thome M., Martillon F., Hofmann K., Rubio V., Steiner V., Schneider P.,  
 RA Mattmann C., Tschopp J.;  
 RT "Equine herpesvirus-2 E10 gene product, but not its cellular  
 RT homologue, activates NF-kappaB transcription factor and c-Jun  
 RT N-terminal kinase.";  
 RL J. Biol. Chem. 274:9962-9968(1999).  
 RN [4]  
 RP SEQUENCE FROM N.A., AND MUTAGENESIS OF LEU-28; LEU-41; ILE-46; LEU-47;  
 RP GLU-53 AND ILE-55.  
 RX MEDLINE=99214590; PubMed=10187815;  
 RA Yan M., Lee J., Schilbach S., Goddard A., Dixit V.M.;  
 RT "E10, a novel caspase recruitment domain-containing proapoptotic  
 RT molecule";  
 RL J. Biol. Chem. 274:10287-10292(1999).  
 RN [5]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=99292766; PubMed=10364242;  
 RA Srinivasula S.M., Ahmed M., Ilin J.-H., Poyet J.-L.,  
 RA Fernandes-Alnemri T., Tsichlis P.N., Alnemri E.S.;  
 RT "CLAP, a novel caspase recruitment domain-containing protein in the  
 RT tumor necrosis factor receptor pathway, regulates NF-kappaB  
 RT activation and apoptosis";  
 RL J. Biol. Chem. 274:17946-17954(1999).  
 RN [6]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Spleen;  
 RX MEDLINE=99329013; PubMed=10400625;  
 RA Costanzo A., Guet C., Vitco P.;  
 RT "C-E10 is a caspase-recruiting domain-containing protein that  
 RT interacts with components of death receptors signaling pathway and  
 RT activates nuclear factor-kappaB";  
 RL J. Biol. Chem. 274:20127-20132(1999).  
 RN [7]  
 RP SEQUENCE FROM N.A., AND VARIANTS MALT LYMPHOMA SER-5; GLU-16; GLU-31;  
 RP ARG-57; LYS-64; GLU-101; PRO-134; ALA-168; SER-174; GLU-213 AND  
 RP ILE-230.  
 RX MEDLINE=99251581; PubMed=10319863;  
 RA Zhang Q., Siebert R., Yan M., Hinzmann B., Cui X., Xue L.,  
 RA Rakestraw K.W., Naeye C.W., Beckmann G., Weisenburger D.D.,  
 RA Sanger W.G., Nowotny H., Vesely M., Callet-Bauchu E., Salles G.,  
 RA Dixit V.M., Rosenthal A., Schlegelberger B., Morris S.W.;  
 RT "Inactivating mutations and overexpression of BCL10, a caspase  
 RT recruitment domain-containing gene, in MALT lymphoma with  
 RT t(1;14)(p22;q32).";  
 RL Nat. Genet. 22:63-68(1999).  
 RN [8]  
 RP PHOSPHORYLATION.  
 RX MEDLINE=21359851; PubMed=11466612;  
 RA Yui D., Yoneda T., Oono K., Katayama T., Imai K., Tshyama M.;  
 RT "Interchangeable binding of Bcl10 to TRAF2 and CIAPs regulates  
 RT apoptosis signaling";  
 RL Oncogene 20:4317-4323(2001).  
 RN [9]  
 RP VARIANTS MESOTHELIOMA SER-5; GLN-45; GLN-58; SER-93; VAL-153; GLU-213





RA Sato N., Gomi F., Morihara T., Mori Y., Miyoshi K., Hitomi U.,  
 RA Ugawa S., Yamada S., Okabe M., Tohyama M.;  
 RT "Regulatory mechanisms of TRAF2-mediated signal transduction by Bcl10,  
 RT a MyD88 lymphoma-associated protein.";  
 RL J. Biol. Chem. 275:11114-11120(2000).  
 CC -1- FUNCTION: Promotes apoptosis, pro-caspase-9 maturation and  
 CC activation of NF-kappaB via NIK and IKK. May be an adapter protein  
 CC between upstream TNFR1-TRADD-RIP complex and the downstream NIK-  
 CC IKK-IKAP complex (By similarity).  
 CC -1- SUBUNIT: Self-associates by CARD-CARD interaction and interacts  
 CC with other CARD-proteins such as CARD9, CARD10, CARD11 and CARD14.  
 CC Binds caspase-9 with its C-terminal domain (By similarity).  
 CC Interacts with TRAF2 and BIRC2/c-IAP2.  
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).  
 CC -1- PTM: Phosphorylated (By similarity).  
 CC -1- SIMILARITY: Contains 1 CARD domain.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL: AB016069; BAA88822.1; -  
 DR InterPro: IPR001315; CARD.  
 DR Pfam: PF00619; CARD; 1.  
 DR SMART: SM00114; CARD; 1.  
 DR PROSITE: PS50209; CARD; 1.  
 KM Apoptosis; Anti-oncogene; Phosphorylation.  
 FT DOMAIN 13 CARD  
 FT SEQUENCE 233 AA; 25999 MW; B43274B4B825FC7D CRC64;  
 SQ  
 Query Match 11.2%; Score 81; DB 1; Length 233;  
 Best Local Similarity 36.5%; Pred. No. 0.63;  
 Matches 31; Conservative 14; Mismatches 32; Indels 8; Gaps 4;  
 QY 13 ALDEBTWEN---MESHRIYRCICPSRLTYPLFOAKYLCOLEBEVLHSERLTNSM 68  
 DB 7 SLNEEDTEVYKDALEMLRYVLCCKILAEHRHFDHLRAKILSRDETEI--SCR-TSSRK 63  
 QY 69 RAGHLDLILKTRGNGAIAFLSEIK 93  
 DB 64 RAGKLDLYQENPK-GIDTLVESIR 87  
 RESULT 10  
 ID VVRC METTH STANDARD; PRT; 579 AA.  
 AC 026541;  
 DT 30-MAY-2000 (Rel. 39, Created)  
 DT 30-MAY-2000 (Rel. 39, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE UVABC System protein C (UVRC protein) (Exonuclease A3C subunit C).  
 GN UVRC OR MTH441.  
 OS Methanobacterium thermoautotrophicum.  
 CC Archaea; Euryarchaeota; Methanobacteria; Methanobacteriales;  
 CC Methanobacteriaceae; Methanochlorobacter.  
 CC NCBI\_TaxID=187420;  
 CX [1]  
 RN SEQUENCE FROM N.A.  
 RC STRAIN=Delta H;  
 RX MEDLINE=98037514; PubMed=9371463;  
 RA Smith D.R., Doucette-Stamm L.A., Delonghery C., Lee H.-M., Dubois J.,  
 RA Aldrege T., Bashirzadeh R., Bickely D., Cook R., Gilbert K.,  
 RA Harrison D., Hoang L., Keagle P., Lamm W., Potcher B., Qiu D.,  
 RA Spadafora R., Vlcare R., Wang Y., Wierzbowski J., Gibson R.,  
 RA Jivan N., Caruso A., Bush D., Safer H., Patwell D., Prabhakar S.,  
 RA McDonnell S., Shimer G., Goyal A., Pietrowski S., Church G.M.,  
 RA Daniels C.J., Mao J.-I., Rice P., Noelling J., Reeve J.N.,  
 RT "Complete genome sequence of Methanobacterium thermoautotrophicum  
 RT delta: functional analysis and comparative genomics.";

RL J. Bacteriol. 179:7135-7155(1997).  
 CC -1- FUNCTION: The UVRC repair system catalyzes the recognition and  
 CC processing of DNA lesions. UVRC both incises the 5' and 3' sides  
 CC of the lesion. The N-terminal half is responsible for the 3'  
 CC incision and the C-terminal half is responsible for the 5'  
 CC incision (By similarity).  
 CC -1- SUBUNIT: Interacts with UVB in an incision complex (By  
 CC similarity).  
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).  
 CC -1- SIMILARITY: Belongs to the uvrc family.  
 CC -1- SIMILARITY: Contains 1 UVR domain.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL: AB000828; AAB84947.1; -  
 DR PIR: F69157; F69157.  
 DR HSRP: P07025; 1B52.  
 DR HAMAP: MF\_00203; -; 1.  
 DR InterPro: IPR000445; HHN.  
 DR InterPro: IPR003583; HHN 1.  
 DR InterPro: IPR001943; UVRB/C.  
 DR InterPro: IPR004791; UVRC.  
 DR InterPro: IPR001162; UVRC\_C.  
 DR InterPro: IPR000305; UVRC\_N.  
 DR Pfam: PF01541; Exci\_endo\_N; 1.  
 DR Pfam: PF00633; HHN; 2.  
 DR Pfam: PF02151; UVR; 1.  
 DR PRODOM: PD008670; UVRC\_C; 1.  
 DR SMART: SM00465; GYC; 1.  
 DR TIGRFAMs: TIGR00194; uvrc; 1.  
 DR PROSITE: PS50151; UVR; 1.  
 DR PROSITE: PS50164; UVRC\_1; 1.  
 DR PROSITE: PS50165; UVRC\_2; 1.  
 KM SOS response; Excision nuclease; DNA repair; DNA recombination;  
 KM DNA excision; Complete proteome.  
 FT DOMAIN 193 228 UVR.  
 FT SEQUENCE 579 AA; 66293 MW; 83D3D7B8F9E3A68 CRC64;  
 SQ  
 Query Match 11.2%; Score 81; DB 1; Length 579;  
 Best Local Similarity 27.6%; Pred. No. 1.9;  
 Matches 43; Conservative 26; Mismatches 53; Indels 34; Gaps 9;  
 QY 7 RDSALTADEETL-----WEMES---HRRIYRCICPSRLTYPLRQ---AVYLQLD 53  
 DB 262 RDGKITGDPIILRGSAPRTIELAFKQYVAIPRVSEILTYPVDDGYIAEWSLR 321  
 QY 54 EEEV-LHSP-----RLTNSAMRAGHLIDLKTRGNGAIAFLSKHNPDDVTLVGL 106  
 DB 322 GEEVYHSPGGAGRLINIAWKASVILKQKAYRDLILQKDKLK--PEIRRMGL 379  
 QY 107 QPDVDFSNFSGSSD-----FDG--LAGTSRNL 134  
 DB 380 -----DISNAGSATGSVAVFIDKPSGSGSYRRYRI 411  
 RESULT 11  
 ID AMPN RABIT STANDARD; PRT; 965 AA.  
 AC P15541;  
 DT 01-APR-1990 (Rel. 14, Created)  
 DT 01-FEB-1996 (Rel. 33, Last sequence update)  
 DT 15-SEP-2003 (Rel. 42, Last annotation update)  
 DE Aminopeptidase N (EC 3.4.11.2) (rBAPN) (Alanyl aminopeptidase)  
 DE (Microsomal aminopeptidase) (Aminopeptidase M) (Leukemia antigen  
 DE CD13).  
 GN ANPEP.

OS Oryctolagus cuniculus (Rabbit).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.  
 NCBI\_TaxID=9986;  
 RN [1]  
 RC SEQUENCE FROM N.A.  
 RX MEDLINE=9409942; Pubmed=7903857;  
 RT "Complete sequence of rabbit kidney aminopeptidase N and mRNA localization in rabbit kidney by in situ hybridization.";  
 RL Biochim. Cell Biol. 71:278-287(1993).  
 RN [2]  
 RP SEQUENCE OF 5-18.  
 RX MEDLINE=82113673; Pubmed=6120002;  
 RA Peracchi H., Maroux S., Bonicel J., Desnuelle P.;  
 RT "The amino acid sequence of the hydrophobic anchor of rabbit intestinal brush border aminopeptidase N";  
 RL Biochim. Biophys. Acta 684:133-136(1982).  
 RN [3]  
 RP SEQUENCE OF 177-965 FROM N.A.  
 RX MEDLINE=90092508; Pubmed=2574692;  
 RA Hansen G.H.; Dabelsteen E., Hoeyer P.E., Olsen J., Sjoestrom H.,  
 RT "Onset of transcription of the aminopeptidase N (leukemia antigen CD 13) gene at the cypc/villus transition zone during rabbit enterocyte differentiation.";  
 RL FEBS Lett. 259:107-112(1989).  
 CC -1- FUNCTION: Broad specificity aminopeptidase. Plays a role in the final digestion of peptides generated from hydrolysis of proteins by gastric and pancreatic proteases. May be involved in the metabolism of regulatory peptides of diverse cell types and in the cleavage of peptides bound to major histocompatibility complex class II molecules of antigen presenting cells. May have a role in angiogenesis (by similarity).  
 CC -1- CATALYTIC ACTIVITY: Release of an N-terminal amino acid, Xaa-|-Xbb-from a peptide, amide or arylamide. Xaa is preferably Ala, but may be most amino acids including Pro (slow action). When a terminal hydrophobic residue is followed by a prolyl residue, the two may be released as an intact Xaa-Pro dipeptide.  
 CC -1- COFACTOR: Binds 1 zinc ion (by similarity).  
 CC -1- SUBUNIT: Homodimer (by similarity).  
 CC -1- SUBCELLULAR LOCATION: Type II membrane protein.  
 CC -1- PFM: Sulfated (by similarity).  
 CC -1- SIMILARITY: Belongs to peptidase family M1.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 CC EMBL: S68687; AAB29334.1; -;  
 DR EMBL: X51508; CA35873.1; -;  
 DR MEROPS: M01.001; -;  
 DR InterPro: IPR001930; Ala\_peptase.  
 DR InterPro: IPR006025; Zn\_MTPeptide.  
 DR Pfam: PF01433; Peptidase\_M1; 1.  
 DR PRINTS: PR00756; ALADIPASE.  
 DR PROSITE: PS00142; ZINC\_PROTEASE; 1.  
 KW Angiogenesis; Hydrolyase; Aminopeptidase; Metalloprotease; Zinc;  
 KW Signal-anchor; Transmembrane; Glycoprotein; Sulfation.  
 FT INIT MET 0  
 FT DOMAIN 1 7 CYTOPLASMIC (POTENTIAL).  
 FT TRANSMEM 8 31 SIGNAL-ANCHOR (TYPE-II MEMBRANE PROTEIN) (POTENTIAL).  
 FT DOMAIN 32 64 CYTOSOLIC SER/THR-RICH JUNCTION.  
 FT DOMAIN 65 965 METALLOPROTEASE.  
 FT METAL 383 383 ZINC (CATALYTIC) (BY SIMILARITY).  
 FT ACT SITE 384 384 BY SIMILARITY.  
 FT METAL 387 387 ZINC (CATALYTIC) (BY SIMILARITY).

FT METAL 406 406 ZINC (CATALYTIC) (BY SIMILARITY).  
 FT ACT SITE 412 472 PROTON DONOR (POTENTIAL).  
 FT MOD\_RES 172 172 SULFATION (POTENTIAL).  
 FT MOD\_RES 414 414 SULFATION (POTENTIAL).  
 FT MOD\_RES 419 419 SULFATION (POTENTIAL).  
 FT MOD\_RES 911 911 SULFATION (POTENTIAL).  
 FT CARBOHYD 39 39 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 124 124 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 258 258 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 314 314 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 551 551 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 569 569 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 623 623 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 733 733 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 816 816 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CONFLICT 9 9 I -> A (IN REF. 2).  
 FT CONFLICT 15 15 I -> F (IN REF. 2).  
 FT CONFLICT 209 209 Q -> QMG (IN REF. 2).  
 FT CONFLICT 225 225 M -> S (IN REF. 2).  
 FT CONFLICT 233 233 P -> L (IN REF. 2).  
 FT CONFLICT 592 593 LE -> QQ (IN REF. 2).  
 SQ SEQUENCE 965 AA; 109186 MW; 012D1895D457A96 CRC64;  
 Query Match 11.1%; Score 80; DB 1; Length 965;  
 Best Local Similarity 28.0%; Pred. No. 4.2;  
 Matches 44; Conservative 20; Mismatches 55; Indels 38; Gaps 10;  
 QY 9 SALTLDEETLWEMSHRIRIVRCIPSR---LTPYL-----RQAKVLCQ 51  
 DB 56 SATTLDDNLRP-----NRVLPKTLIDSYNVVLRPLTSNSGLVFTGSSIVRFICQ 110  
 QY 52 -LDSEEVHSPRLTNSARAGHLLDLTRGKN-GAIAFELSLEKHPNDVY---TLVTG 105  
 DB 111 EATVITIIHSKLL-NYITIGHRVVLRGVRSQPPALISTELVELTVVHLQGLVAG 169  
 QY 106 LQPDVDFSNFSGESDFFGLAGTSR-----NRLILV 136  
 DB 170 SQYEND-TQFQGLLA--DDLGFYRSEYMEGNVRKV 203  
 RESULT 12  
 KC64 HUMAN STANDARD; PRT; 519 AA.  
 ID KC64 HUMAN  
 AC O8DN1; Q96H24;  
 DT 28-FEB-2003 (Rel. 41, Created)  
 DT 28-FEB-2003 (Rel. 41, Last sequence update)  
 DT 15-SEP-2003 (Rel. 42, Last annotation update)  
 DE Potassium voltage-gated channel subfamily G member 4 (Potassium channel Kv6.3).  
 GN KCNG4 OR KCNG3.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
 NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A. (ISOFORM 1).  
 RC TISSUE=Brain;  
 RX MEDLINE=22388257; Pubmed=12477932;  
 RC TISSUE=Muscle;  
 RX MEDLINE=22388257; Pubmed=12477932;  
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,  
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bat N.K.,  
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,  
 RA Diatchenko L., Murusani K., Farmer A.A., Rubin G.M., Hong L.,  
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,  
 RA Brownstein W.J., Uediri T.B., Toshlyuk S., Carninci P., Prange C.,

RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,  
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
 RA Richards S., Morley L.J., Hale S.C., Garcia A.M., Gay L.J., Hulyk S.W.,  
 RA Villalón D.K., Murry D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
 RA Fahy J., Helton E., Ketterman M., Maden A., Rodrigues S., Sanchez A.,  
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
 RA Blakesley R.W., Touchman J.W., Green B.D., Dickson M.C.,  
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,  
 RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,  
 RA Scherch A., Schein J.E., Jones S.J.M., Marra M.A.,  
 RA "Generation and initial analysis of more than 15,000 full-length  
 RT human and mouse cDNA sequences.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).  
 CC -1- FUNCTION: Potassium channel subunit. Modulates channel activity by  
 CC shifting the threshold and the half-maximal activation to more  
 CC negative values.  
 CC -1- SUBUNIT: Heteromultimer with KCNB1, KCNC1 and KCNF1. Does not form  
 CC homomultimers.  
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein. Has to be  
 CC associated with KCNB1 or possibly another partner to get inserted  
 CC in the plasma membrane. Remains intracellular in the absence of  
 CC KCNB1.  
 CC -1- ALTERNATIVE PRODUCTS:  
 CC Event=Alternative splicing; Named isoforms=2;  
 CC Name=1;  
 CC IsoId=Q8TDN1-1; Sequence=Displayed;  
 CC Name=2;  
 CC IsoId=Q8TDN1-2; Sequence=VSP\_001029, VSP\_001030;  
 CC Note=No experimental confirmation available;  
 CC -1- TISSUE SPECIFICITY: Highly expressed in brain, and at lower levels  
 CC in liver, small intestine and colon.  
 CC -1- DOMAIN: The segment S4 is probably the voltage-sensor and is  
 CC characterized by a series of positively charged amino acids at  
 CC every third position.  
 CC -1- SIMILARITY: BELONGS TO THE VOLTAGE-GATED POTASSIUM CHANNEL FAMILY.  
 CC G SUBFAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL: AF348984; AAL83911.1; -;  
 DR EMBL: BC008969; AAH08969.1; -;  
 DR Genew: HGNC:19697; KCNG4.  
 DR MIM: 607603; -;  
 DR InterPro: IPR000210; BTB\_POZ.  
 DR InterPro: IPR005821; Ion\_trans.  
 DR InterPro: IPR001622; K+channel\_pore.  
 DR InterPro: IPR003131; K\_tetra.  
 DR InterPro: IPR005820; M+channel\_nlg.  
 DR Pfam: PF00520; Ion\_trans; 1.  
 DR Pfam: PF02214; K\_tetra; 1.  
 DR SMART: SMO0225; BTB; 1.  
 DR Transport: Ion transport; Ionic channel; Voltage-gated channel;  
 KM Potassium channel; Potassium transport; Potassium; Transmembrane;  
 KM Multigene family; Alternative splicing.  
 KM DOMAIN 1 118 CYTOPLASMIC (POTENTIAL).  
 FT TRANSMEM 119 139 SEGMENT S1 (POTENTIAL).  
 FT TRANSMEM 259 279 SEGMENT S2 (POTENTIAL).  
 FT DOMAIN 280 289 CYTOPLASMIC (POTENTIAL).  
 FT TRANSMEM 290 310 SEGMENT S3 (POTENTIAL).  
 FT TRANSMEM 335 355 SEGMENT S4 (POTENTIAL).  
 FT TRANSMEM 356 370 CYTOPLASMIC (POTENTIAL).  
 FT TRANSMEM 371 391 SEGMENT S5 (POTENTIAL).  
 FT DOMAIN 406 426 SEGMENT H5 (POTENTIAL).  
 FT TRANSMEM 434 454 SEGMENT S6 (POTENTIAL).  
 FT DOMAIN 455 519 CYTOPLASMIC (POTENTIAL).  
 FT TRANSMEM 378 378 POLY-LEU.  
 FT VARSPIC 253 GECS -> VSGL (in isoform 2).

FT FT /FTId=VSP\_001029.  
 FT VARSPIC 257 519 Missing (in isoform 2).  
 FT FT /FTId=VSP\_001030.  
 SQ SEQUENCE 519 AA; 58979 MW; 55BBA355931AB0A4 CRC64;  
 Query Match 11.0%; Score 79; DB 1; Length 519;  
 Best Local Similarity 25.8%; Pred. No. 2.6;  
 Matches 33; Conservative 21; Mismatches 56; Indels 18; Gaps 5;  
 QY 3 ELCDSDALRALDDELTEMWESHRHIVACIPSRITPTLRQAKYLCQIDEEVYHSR 62  
 DB 137 EMC---ALSPQELAYWGIEBAH---LEEC-CIRKULRELEBEIAKURDVLKQR 188  
 QY 63 LT---NSAMRAGHLDLILKTRGN-----GAIAPLESIKFNPDPVYTLVTGLQDPVDF 112  
 DB 189 ETRPAPSHSRWIGCMKRLKEMENPQSGAPGVAFALSLTFATYAVSLCVSTMPDLRA 248  
 QY 113 SNFSGESS 120  
 DB 249 EDDGEGCS 256  
 RESULT 13  
 ID BCLA MOUSE STANDARD; PRT; 233 AA.  
 AC 0920H7;  
 DT 28-FEB-2003 (Rel. 41, Created)  
 DT 28-FEB-2003 (Rel. 41, Last sequence update)  
 DT 15-SEP-2003 (Rel. 42, Last annotation update)  
 DE B-cell lymphoma/leukemia 10 (B-cell CLL/lymphoma 10) (Bcl-10) (CED-3/ICH-1 prodomain homologous E10-like regulator) (mCIPER) (CARD-3(CARMEN) (Mammalian CARD-containing adapter molecule E10) (mE10) (Cellular-E10) (C-E10) (CARD-like apoptotic protein) (mCLAP).  
 DE Bcl10 OR CIPER OR CLAP.  
 GN Bcl10 OR CIPER OR CLAP.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A. PubMed=9989495;  
 RX MEDLINE=99142601; PubMed=9989495;  
 RA Willis T.G., Jadayel D.M., Du M.-Q., Peng H., Perry A.R.,  
 RA Abdul-Rauq M., Price H., Kairan L., Mayekodunni O., Wlodarska I.,  
 RA Pan L., Crook T., Hamoudi R., Isaacson P., Dyer M.J.S.;  
 RT "Bcl10 is involved in t(1;14)(p22;q32) of MALT B cell lymphoma and  
 RT mutated in multiple tumor types.";  
 RL Cell 96:35-45(1999).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=99214545; PubMed=10187770;  
 RA Kosaki T., Inohara N., Chen S., Carrioe R., Merino J., Hottinger M.O.,  
 RA Nabel G.J., Nunez G.;  
 RT "CIPER, a novel NF-kappaB-activating protein containing a caspase  
 RT recruitment domain with homology to Herpesvirus-2 protein E10.";  
 RL J. Biol. Chem. 274:9955-9961(1999).  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=99214546; PubMed=10187771;  
 RA Thome M., Martillon P., Hofmann K., Rubio V., Steiner V., Schneider P.,  
 RA Matman C., Tschopp J.;  
 RT "Equine herpesvirus-2 E10 gene product, but not its cellular  
 RT homologue, activates NF-kappaB transcription factor and c-Jun N-  
 RT terminal kinase.";  
 RL J. Biol. Chem. 274:9962-9968(1999).  
 RN [4]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=99214590; PubMed=10187815;  
 RA Yan M., Lee J., Schilbach S., Goddard A., Dixit V.M.;  
 RT "ME10, a novel caspase recruitment domain-containing proapoptotic  
 RT molecule.";  
 RL J. Biol. Chem. 274:10287-10292(1999).

RN [5]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=92922766; PubMed=10364242;  
 RA Scrinivasula S.M., Ahmad M., Lin J.-H., Poyet J.-L.,  
 RA Fernandez-Alnemati T., Tschilis P.N., Alnemati E.S.,  
 RT "ClAP, a novel caspase recruitment domain-containing protein in the  
 RT tumor necrosis factor receptor pathway, regulates NF-kappaB activation  
 RT and apoptosis.";  
 RL J. Biol. Chem. 274:17946-17954 (1999).  
 RN [6]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=22388257; PubMed=12477932;  
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,  
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,  
 RA Diatchenko L., Marusik K., Farmer A.A., Rubin G.M., Hong L.,  
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,  
 RA Brownstein M.J., Udell T.B., Toshiyuki S., Carninci P., Prange C.,  
 RA Rana S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaney S.J.,  
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
 RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
 RA Fahey J., Helton E., Kettman W., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
 RA Blakeley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,  
 RA Butcherfield Y.S.N., Krzywinski M.I., Skalska U., Smalins D.E.,  
 RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.,  
 RT "Generation and initial analysis of more than 15,000 full-length  
 RT human and mouse cDNA sequences.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).  
 CC -1- FUNCTION: Promotes apoptosis, pro-caspase-9 maturation and  
 CC activation of NF-kappaB via NIK and IKK. May be an adapter protein  
 CC between upstream TNFR1-TRADD-RIP complex and the downstream NIK-  
 CC IKK-IKAP complex (By similarity).  
 CC -1- SUBUNIT: Self-associates by CARD-CARD interaction and interacts  
 CC with other CARD-proteins such as CARD9, CARD10, CARD11 and CARD14.  
 CC Binds caspase-9 with its C-terminal domain. Interacts with TRAF2  
 CC and BIRC2/C-IAP2 (By similarity).  
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).  
 CC -1- TISSUE SPECIFICITY: Highly expressed in heart, brain, spleen,  
 CC lung, liver, skeletal muscle, kidney and testis. Detected in  
 CC developing brain, olfactory epithelium, tongue, whisker follicles,  
 CC salivary gland, heart, lung, liver and intestinal epithelia of  
 CC stage 15 embryos.  
 CC -1- PTM: Phosphorylated (By similarity).  
 CC -1- SIMILARITY: Contains 1 CARD domain.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 CC -----  
 CC EMBL: AJ006289; CAA06956.1; -  
 CC EMBL: AF057701; AAD15801.1; -  
 CC EMBL: AF100339; AAD16429.1; -  
 CC EMBL: AF127387; AAD32598.1; -  
 CC EMBL: AF134396; AAD39148.1; -  
 CC EMBL: BC024379; AAH24379.1; -  
 CC MGD: MGI:1337994; Bcl10.  
 CC GO: GO:0006915; P:apoptosis; IDA.  
 CC GO: GO:0016066; P:humoral defense response (sensu Vertebrata); IMP.  
 CC GO: GO:0007249; P:NIK-I-kappaB/NF-kappaB cascade; IMP.  
 CC InterPro: IPR001315; CARD.  
 CC Pfam: PF00619; CARD; 1.  
 CC SMART: SMO0114; CARD; 1.  
 CC PROSITE: PSS0209; CARD; 1.  
 CC Apoptosis; Anti-oncogene; Phosphorylation.

FT DOMAIN 13 101 CARD  
 SQ SEQUENCE 233 AA; 25948 MW; C0539BC97102DB8 CRC64;  
 Query Match 10.8%; Score 78; DB 1; Length 233;  
 Best Local Similarity 35.3%; Pred. No. 1.3;  
 Matches 30; Conservative 14; Mismatches 33; Indels 8; Gaps 4;  
 QY 13 ALDEETLWEM---MESHRRHIVRCICPSRLTPYIRQAKVLCQDEEVLHSPRLTNSAM 68  
 DB 7 SLTEEDLVEKKDALENIRVYLCEKIIAERFHDHRAKXILSRDETEI--SCR-TSSRX 63  
 QY 69 RAGHLDLTKRGKNGALAFLESK 93  
 DB 64 RAGKLDLYIQ-ENPRGLDTVESIR 87  
 RESULT 14  
 ID STYL\_HUMAN STANDARD; PRT; 685 AA.  
 AC Q13586;  
 DT 28-FEB-2003 (Rel. 41, Created)  
 DT 28-FEB-2003 (Rel. 41, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE Stromal interaction molecule 1 precursor.  
 GN STIM1 OR GOK.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxId=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Fetal liver, and Placenta;  
 RX MEDLINE=97079692; PubMed=8921403;  
 RA Parker N.U., Begley C.G., Smith P.J., Fox R.M.;  
 RT "Molecular cloning of a novel human gene (D11S489E) at chromosomal  
 RT region 11p15.5.";  
 RL Genomics 37:253-256 (1996).  
 RN [2]  
 RP GLYCOSYLATION, PHOSPHORYLATION, SUBCELLULAR LOCATION, AND TISSUE  
 RP SPECIFICITY.  
 RX MEDLINE=20461006; PubMed=11004585;  
 RA Manji S.S., Parker N.J., Williams R.T., van Stekelenburg L.,  
 RA Pearson R.B., Dziadek M., Smith P.J.;  
 RT "STIM1: a novel phosphoprotein located at the cell surface.";  
 RL Biochem. Biophys. Acta 1481:147-155 (2000).  
 RN [3]  
 RP TISSUE SPECIFICITY, AND SUBUNIT.  
 RX MEDLINE=21356314; PubMed=11463338;  
 RA Williams R.T., Manji S.S.M., Parker N.J., Hancock M.S.,  
 RA Van Stekelenburg L., Eid U.-P., Senior P.V., Kazemwadel J.S.,  
 RA Shandaula T., Saito R., Smith P.J., Dziadek M.A.;  
 RT "Identification and characterization of the STIM (stromal interaction  
 RT molecule) gene family: coding for a novel class of transmembrane  
 RT proteins.";  
 RL Biochem. J. 357:673-685 (2001).  
 RN [4]  
 RP DISEASE.  
 RX MEDLINE=98021968; PubMed=9377559;  
 RA Sabloni S., Barbanti-Brodano G.,  
 RT "GOK: a gene at 11p15 involved in rhabdomyosarcoma and rhabdoid tumor  
 RT development.";  
 RL Cancer Res. 57:4493-4497 (1997).  
 CC -1- FUNCTION: Possible adhesion molecule with a role in early  
 CC hematopoiesis by mediating attachment to stromal cells. Influences  
 CC the survival and/or proliferation of B cell precursors. Binding to  
 CC cells requires Mn(II) (By similarity).  
 CC -1- SUBUNIT: Oligomer with STIM2.  
 CC -1- SUBCELLULAR LOCATION: Type I membrane protein. Cell surface  
 CC (Potential).  
 CC -1- TISSUE SPECIFICITY: Ubiquitously expressed in various human  
 CC primary cells and tumor cell lines.  
 CC -1- PTM: Glycosylated.  
 CC -1- PTM: Phosphorylated predominantly on Ser residues.

```

CC -1- DISEASE: Defects in STM1 may cause rhabdomyosarcoma and rhabdoid
CC tumors.
CC -1- SIMILARITY: Contains 1 sterile alpha motif (SAM) domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: U52426; AAC51627.1; -.
CC GenBank: HGNC:11386; STM1.
CC MIM: 605921; -.
CC DR GO:000587; C:integral to plasma membrane; TAS.
CC DR GO:0005194; P:cell adhesion molecule activity; NMS.
CC DR GO:0005515; P:protein binding activity; IPI.
CC DR GO:0001637; P:cell-cell adhesion; NMS.
CC DR GO:0008284; P:positive regulation of cell proliferation; TAS.
CC DR InterPro: IPR001660; SAM.
CC DR SMART: SM00454; SAM; 1.
CC DR PROSITE: PS50105; SAM DOMAIN; 1.
CC DR Cell adhesion; Anti-oncogene; Transmembrane; Coiled coil; Signal;
CC Glycoprotein; Phosphorylation.
CC FT SIGNAL 1 22 POTENTIAL.
CC FT CHAIN 23 685 STROMAL INTERACTION MOLECULE 1.
CC FT DOMAIN 23 213 EXTRACELLULAR (POTENTIAL).
CC FT TRANSMEM 214 234 POTENTIAL.
CC FT DOMAIN 235 685 CYTOPLASMIC (POTENTIAL).
CC FT DOMAIN 132 200 SAM.
CC FT DOMAIN 238 343 COILED COIL (POTENTIAL).
CC FT DOMAIN 270 336 GLU-RICH.
CC FT DOMAIN 362 390 COILED COIL (POTENTIAL).
CC FT DOMAIN 600 629 PRO-SER-RICH.
CC FT DOMAIN 672 685 IYS-RICH.
CC FT CARBOHYD 131 131 N-LINKED (GLCNAC. . .) (POTENTIAL).
CC FT CARBOHYD 171 171 N-LINKED (GLCNAC. . .) (POTENTIAL).
CC FT CARBOHYD 171 171 N-LINKED (GLCNAC. . .) (POTENTIAL).
CC SQ SEQUENCE 685 AA; 77492 MW; 0AB512CA8D68A7A7 CRC64;

Query Match 10.5%; Score 75.5; DB 1; Length 685;
Best Local Similarity 24.6%; Pred. No. 8;
Matches 31; Conservative 14; Mismatches 48; Indels 33; Gaps 4;

QY 16 EETLWEMESHRRIVRCICPSRLTPYLQAKVLCQDDEEVLHSPRLTNSAMRA----- 70
DB 545 DEALNMTNGRHRLIEGVHPSLV-----EKLPSPALAKYALALNHL 590
QY 71 --GHLLDLKTRGKNGAIAFLSLKFNPDVYTLVGLQPDVDFSNFSGSSDPDLAQT 128
DB 591 DKASHLMELSPSPAPGSPHLDSSRSHSPS-----SPDPTSPVGDSSALQ-----A 638
QY 129 SRNRL 134
DB 639 SRNTRI 644

RESULT 15
YMSO_CABEL STANDARD; PRT; 290 AA.
AC Q21268;
DT 15-JUL-1998 (Rel. 36, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Hypothetical protein K07C11.7 in chromosome V.
GN K07C11.7
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae;
OC Rhabditidae; Pelodierinae; Caenorhabditis.
OX NCBI_Taxid=6239;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;

```

```

RA Wu X., Le T.T.;
RA Submitted (APR-1996) to the EMBL/GenBank/DBJ databases.
RN [2]
RP REVISIONS.
RA Waterston R.;
RL Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: BELONGS TO THE UPF0046 FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: U53336; AAA96177.2; -.
CC DR WormPep: K07C11.7; CE28592.
CC DR InterPro: IPR004843; M-peptidase.
CC DR Pfam: PF00149; Metallophos; 1.
CC DR Hypothetical protein.
CC SQ SEQUENCE 290 AA; 32905 MW; F18ED90B2B8949A2 CRC64;

Query Match 10.4%; Score 75; DB 1; Length 290;
Best Local Similarity 26.5%; Pred. No. 3.3;
Matches 36; Conservative 19; Mismatches 61; Indels 20; Gaps 6;

QY 14 LDEETLWEMESHRRIVRCIC-----PSRLTPYLQAKVLCQDDEEVLHSPR 62
DB 27 IDPDAENELMDSIKHTVQIVKNAIGELNSPTDGTPL---KYVCISDTHEQLHNT 83
QY 63 LTNAMRAGHLIDLKTRGNGAIAFLSLKFNPDVYTLVGLQPDVDFSN--FSGS 119
DB 84 VPDGDVLI-HAGDFTNNGKEELIKFVEMTRF--PHKYLVVAGNHELGFDDENQGER 140
QY 120 SDFGLAGTSRNRL 135
DB 141 QADKXGIGTGDGYNIL 156

Search completed: February 18, 2004, 04:45:08
Job time : 28 secs

```

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

CM protein - protein search, using sw model

Run on: February 18, 2004, 04:30:32 ; Search time 85 Seconds

(without alignments)  
421.992 Million cell updates/sec

Title: US-10-032-159A-16

Perfect score: 720  
Sequence: 1 MGELCRDRLTALDEBTLW.....SDFDGLAGTSRNRLIVTX 139

Scoring table: BLOSUM62  
Gapop 10.0 , Gapex 0.5

Searched: 830525 seqs, 258052604 residues

Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

SPTREMBL\_23:\*  
1: sp\_archaea:\*  
2: sp\_bacteria:\*  
3: sp\_fungi:\*  
4: sp\_human:\*  
5: sp\_invertebrate:\*  
6: sp\_mammal:\*  
7: sp\_mhc:\*  
8: sp\_organelle:\*  
9: sp\_phage:\*  
10: sp\_plant:\*  
11: sp\_protist:\*  
12: sp\_virus:\*  
13: sp\_vertebrate:\*  
14: sp\_unclassified:\*  
15: sp\_virus:\*  
16: sp\_bacteriap:\*  
17: sp\_archaeap:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	270.5	37.6	1159	11	Q8CIS0 mus musculi
2	270.5	37.6	1171	4	Q8TES3 homo sapien
3	215.5	29.9	535	13	Q8JFQ3 cyprinus ca
4	80	11.1	2114	10	Q9C6Y4 arabidopsi
5	77.5	10.8	2548	5	Q8I895 plasmodium
6	76.5	10.6	832	16	Q914C4 pseudomonas
7	75.5	10.5	587	12	Q9YV93 macropodid
8	74.5	10.3	394	2	Q9F5M0 pseudomonas
9	74	10.3	210	12	Q66877 equine herp
10	74	10.3	311	12	Q9YVNS equine herp
11	73.5	10.2	315	5	Q9NA19 fuscobacteri
12	73.5	10.2	382	16	Q8R6B3 fuscobacteri
13	73	10.1	265	16	Q97EL5 clostridium
14	73	10.1	447	16	Q8CV98 oceanobacill
15	72.5	10.1	295	5	Q76407 caenorhabdi
16	72.5	10.1	701	16	Q8CST1 staphylococ

17	72.5	10.1	845	5	Q9Y466 drosophila
18	72.5	10.1	909	10	Q9M5A1 arabidopsis
19	72.5	10.1	1318	5	Q19733 caenorhabdi
20	72	10.0	209	4	Q8NI97 homo sapien
21	72	10.0	230	17	Q8TRF2 methanosarc
22	72	10.0	664	10	Q9S714 arabidopsis
23	72	10.0	1234	5	Q24690 drosophila
24	71.5	9.9	660	10	Q9SNE4 arabidopsis
25	71.5	9.9	707	4	Q9NVK3 homo sapien
26	71.5	9.9	894	12	Q9YV92 macropodid
27	71.5	9.9	956	4	Q9NV73 homo sapien
28	71.5	9.9	956	4	Q8NEC0 homo sapien
29	71.5	9.9	1141	4	Q9EAE7 homo sapien
30	71.5	9.9	2099	4	Q9VNU8 homo sapien
31	71.5	9.9	2099	4	Q9Y2W9 homo sapien
32	71	9.9	213	17	Q9YVZ2 sulfobacill
33	71	9.9	265	5	Q61218 caenorhabdi
34	71	9.9	310	16	Q8R888 thermococci
35	71	9.9	379	17	Q8T711 methanosarc
36	71	9.9	428	16	Q8ZF14 yersinia pe
37	70.5	9.8	263	11	Q60407 cricetus cr
38	70.5	9.8	329	16	Q916H2 pseudomonas
39	70.5	9.8	735	10	Q91N85 arabidopsis
40	70.5	9.8	2215	10	Q8LH24 oryza sativ
41	70	9.7	305	17	Q976C0 sulfobacill
42	70	9.7	308	10	Q8LPM3 helianthus
43	70	9.7	525	12	Q9MMJ0 measles vir
44	70	9.7	599	16	Q8YH20 bruceella me
45	70	9.7	741	12	Q9YIS6 xestia c-ni

## ALIGNMENTS

RESULT 1	Q8CIS0	PRELIMINARY;	PRT; 1159 AA.
ID	Q8CIS0		
AC	Q8CIS0		
DT	01-MAR-2003 (TREMBLrel. 23, Created)		
DE	01-MAR-2003 (TREMBLrel. 23, Last sequence update)		
DE	01-MAR-2003 (TREMBLrel. 23, Last annotation update)		
DR	Caspase recruitment domain family member 11.		
GN	CARD11.		
OS	Mus musculus (Mouse).		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.		
OX	NCBI_TaxID=10090;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RC	TISSUE=Thymus;		
RA	MEDLINE=2244060; PubMed=12356734;		
RX	Pomerantz J.L., Denny E.M., Baltimore D.;		
RT	"CARD11 mediates factor-specific activation of NF-kappaB by the T cell		
RT	receptor complex";		
RU	EMBO J. 21:5184-5194(2002).		
DR	EMBL; AY135367; AANIO150.1; -		
SQ	SEQUENCE 1159 AA; 134039 MW; 8947A3CD5D0D81CD CRC64;		
Query Match	37.6%; Score 270.5; DB 11; Length 1159;		
Best Local Similarity	48.6%; Pred. No. 1.3e-19;		
Matches	54; Conservative 23; Mismatches 33; Indels 1; Gaps 1;		
QY	15 DEETLWEMSHRRHIVRCICPSRLTPYLRQAKVLCOLDEEVLHSPRLTNSAMRAGHLL 74		
DB	18 EEELALMDNVECNHRHLSRYINPAKLTPLYRQCKVIDEDQDEDEVLANAPMLPSKINAGRL 77		
QY	75 DLILTRGNKAIAFLSELSKFNPNPDVYTVLTGQLQPDVDSNFSGSSDDPGL 125		
DB	78 DILHTGQGVVFLSELSFYPPELYKLVTKGKPTFRFSTIVVEBG-HEGL 127		
RESULT 2	Q8TES3		

```

ID 08TES3 PRELIMINARY; PRT; 1171 AA.
AC 08TES3;
DT 01-JUN-2002 (TREMBLrel. 21, Created)
DT 01-JUN-2002 (TREMBLrel. 21, Last sequence update)
DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)
DE FLJ00120 protein (Fragment).
GN FLJ00120.
OS Homo sapiens (human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxId=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Spleen;
RA Ohara O., Nagase T., Kikuno R., Okumura K.;
RT "The nucleotide sequence of a long cDNA clone isolated from human
RT spleen."
RL Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AK074049; BAB84875.1; -
DR InterPro; IPR001315; CARD.
DR SMART; SM00228; PDZ; 1.
DR PROSITE; PS50209; CARD; 1.
FT NON_TER
SQ SEQUENCE 1171 AA; 134966 MW; FA567ABBC8A703FF CRC64;

Query Match 37.6%; Score 270.5; DB 4; Length 1171;
Best Local Similarity 48.6%; Pred. No. 1.3e-19;
Matches 54; Conservative 23; Mismatches 33; Indels 1; Gaps 1;

QY 15 DEEFLWEMESHRRIVRCICPSRLTPYLRQAKVLCOLDEEYVLSHPRITNSAMRAGHLL 74
DB 35 EEDALWENECNRMHLSRYINPAKLTPIYLRQCKVIDQEDDEVLANPMLPSKINRAGRL 94

QY 75 DLKTRGKNGAIAFLSLKFNHPDVTYLTGLOPDVDFSNFSGSSDPDGL 125
DB 95 DILHTGGRGVVFLSLERYVELKVLTKGKPTFRSTIVVSG-HBGL 144

RESULT 3
Q8F03 PRELIMINARY; PRT; 535 AA.
AC 08F03;
DT 01-OCT-2002 (TREMBLrel. 22, Created)
DT 01-OCT-2002 (TREMBLrel. 22, Last sequence update)
DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)
DE Caspase recruitment domain protein 9.
GN CARD9.
OS Cyprinus carpio (Common carp).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC Cyprinidae; Cyprinus.
OX NCBI_TaxId=7962;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Head kidney;
RA Kono T., Sakai M.;
RT "Molecular cloning of a novel caspase recruitment domain protein 9
RT (CARD9) in teleosts."
RL Submitted (JUN-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB086061; BAC0527.1; -
DR InterPro; IPR001315; CARD.
DR PROSITE; PS50209; CARD; 1.
SQ SEQUENCE 535 AA; 63375 MW; 9F3265D06C452D4D CRC64;

Query Match 29.9%; Score 215.5; DB 13; Length 535;
Best Local Similarity 38.2%; Pred. No. 3.3e-14;
Matches 52; Conservative 21; Mismatches 50; Indels 13; Gaps 2;

QY 15 DEEFLWEMESHRRIVRCICPSRLTPYLRQAKVLCOLDEEYVLSHPRITNSAMRAGHLL 74
DB 12 EDEECARLDEYRLMLIKTIEPSRIITYLRQCKVLSSEDEQLYNPDSLVIRRRKYGML 71

```

```

QY 75 DLKTRGKNGAIAFLSLKFNHPDVTYLTGLOPDVDFSNFSGSSDPDGL 121
DB 72 DILRTGKNGYEALFSLIEDLPDVTYKRTGKEPARVSVLIDYAGCGLOFLMSEVSR 131

QY 122 FDGLAGTSRNLRLVLT 137
DB 132 LQKLAQDERRRKRLV 147

RESULT 4
Q9C6Y4 PRELIMINARY; PRT; 2114 AA.
AC 09C6Y4;
DT 01-JUN-2001 (TREMBLrel. 17, Created)
DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)
DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)
DE Hypothetical 230.9 kDa protein.
GN T7023.25.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eustosida II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxId=3702;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Columbia;
RX MEDLINE=21016719; PubMed=11130712;
RA Theologis A., Ecker J.R., Palm C.J., Federspiel N.A., Kaul S.,
RA White O., Alonso J., Altieri H., Araujo R., Bowman C.L., Brooks S.Y.,
RA Buehlerr E., Chan A., Chao Q., Chen H., Cheuk R.F., Chin C.W.,
RA Chung M.K., Conn L., Conway A.B., Conway A.R., Creasy T.H., Dewar K.,
RA Dunn P., Egu P., Feldblum T.V., Feng J.-D., Fong B., Fujii C.Y.,
RA Gill J.E., Goldsmith A.D., Haas B., Hansen N.F., Hughes B., Huzar L.,
RA Hunter J.V., Jenkins J., Johnson-Hopson C., Khan S., Khaykin E.,
RA Kim C.J., Koo H.L., Kremenetskaia I., Kurtz D.B., Kwan A., Lam B.,
RA Langin-Hooper S., Lee A., Lee J.M., Lenz C.A., Li J.H., Li Y.-P.,
RA Lin X., Liu S.X., Liu Z.A., Lurco J.S., Malt R., Marzilli A.,
RA Miltseher J., Miranda M., Nguyen M., Nieman W.C., Osborne B.I.,
RA Pat G., Peterson J., Pham P.K., Rizzo M., Rooney T., Rowley D.,
RA Sakano H., Salzberg S.L., Schwartz J.R., Shinn P., Southwick A.M.,
RA Sun H., Tallon L.J., Tambunga G., Tarrum M.J., Town C.D.,
RA Uterback T., Van Aken S., Vaysberg M., Vysotskaia V.S., Walker M.,
RA Wu D., Yu G., Fraser C.M., Venter J.C., Davis R.W.;
RT "Sequence and analysis of chromosome 1 of the plant Arabidopsis
RT thaliana."
RL Nature 408:816-820 (2000).
DR EMBL; AC074228; AAC50555.1; -
DR InterPro; IPR000225; Armadillo.
DR InterPro; IPR000008; C2.
DR Pfam; PF00514; Armadillo_egg; 13.
DR Pfam; PF00514; Armadillo_egg; 13.
DR Pfam; PF00514; Armadillo_egg; 13.
DR SMART; SM00168; C2; 1.
DR SMART; SM00168; ARM; 9.
DR SMART; SM00239; C2; 1.
DR SMART; SM00239; C2; 1.
KV Hypothetical protein.
SQ SEQUENCE 2114 AA; 230851 MW; 76E0C52A6B909AC4 CRC64;

Query Match 11.1%; Score 80; DB 10; Length 2114;
Best Local Similarity 26.5%; Pred. No. 40;
Matches 40; Conservative 22; Mismatches 69; Indels 20; Gaps 5;

QY 4 LCRDPSALTALDE--ETLWEMESHRR-----IVRCISRLTPYLRQAKVLCOL 52
DB 1206 LSPQDSTEITVSELSLFRSPETIRHNTAISNKOIGIHLSRSRY-NAARVLCET 1264

QY 53 DEEFLWEMESHRRIVRCICPSRLTPYLRQAKVLCOLDEEYVLSHPRITNSAMRAGHLL 106
DB 1265 FSEHIRPSELAWKRL--SLIEMWNTLSEERYAALTAVALYKLMGINPRDIITSLEGN 1322

QY 107 QPDVDFSNFSGSSDPDGLAGTSRNLRLVLT 137
DB 1323 PLDNIYKTLISLSSLSSEKTSARICRFLT 1353

```

## RESULT 5

Q81E95 ID 081E95 PRELIMINARY; PRT; 2548 AA.  
 AC 081E95;  
 DT 01-MAR-2003 (TRENBLREL. 23, Created)  
 DT 01-MAR-2003 (TRENBLREL. 23, Last sequence update)  
 DT 01-MAR-2003 (TRENBLREL. 23, Last annotation update)  
 DE Hypothetical protein.  
 GN M113P1.122.  
 OS Plasmodium falciparum (isolate 3D7).  
 OC Eukaryota; Alveolata; Apicomplexa; Haemosporidia; Plasmodium.  
 OX NCBI\_TaxID=36329;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Harris B., Lennard N., Clark L., Line A., Barron A., Corton C.,  
 RA Berriman M., Pain A., Hall N., Atkin R., Chillingworth C., Doggett J.,  
 RA Omond D., Sanders M., Hayes R., Hall S., Quail M., Barrell B.,  
 RL Submitted (SFP-2002) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AL844509; CAD52368.1; -  
 KM Hypothetical protein.  
 SQ SEQUENCE 2548 AA; 300711 MW; A0F2BC3015CC4AE CRC64;

Query Match 10.8%; Score 77.5; DB 5; Length 2548;  
 Best Local Similarity 22.5%; Pred. No. 93;  
 Matches 27; Conservative 21; Mismatches 53; Indels 19; Gaps 3;

QY 25 SHRRIRYRCIC-----PSRLTPYRQAKVLCQDDEEVLSPLRTNSAMRAGHLLDL 76  
 DB 1191 SHKRIRICANHMVDYKKEHMYLKYIKVYDICKDEP--IHNNNNN-----NI 1239  
 QY 77 LKTRGKNGAIAFLSKFKHPNPDVYTLVGIQPDVFNFGSSDPDGLAGTSRNRLLY 136  
 DB 1240 LSEKGYGDENNLSISFKHNNSYEPKXKCTCKGDSNYLSHDNKGITDMNNNTKMY 1299

## RESULT 6

Q91A4 ID 091A4 PRELIMINARY; PRT; 532 AA.  
 AC 091A4;  
 DT 01-MAR-2001 (TRENBLREL. 16, Created)  
 DT 01-MAR-2001 (TRENBLREL. 16, Last sequence update)  
 DT 01-MAR-2003 (TRENBLREL. 23, Last annotation update)  
 DE Hypothetical protein PA1214.  
 GN PA1214.  
 OS Pseudomonas aeruginosa.  
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;  
 OC Pseudomonadaceae; Pseudomonas.  
 OX NCBI\_TaxID=287;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA STRAIN=ATCC 15692 / PA01;  
 RX MEDLINE=20437337; PubMed=10984043;  
 RA Stever C.K., Pham X.-Q.T., Erwin A.L., Mizoguchi S.D., Warener P.,  
 RA Hickey M.J., Brinkman F.S.L., Hutnagle W.O., Kowalik D.J., Lagrou M.,  
 RA Gader R.L., Goltzy L., Tolentini E., Westbrock-Wadman S., Yuan Y.,  
 RA Brody L.L., Coulter S.N., Folger K.R., Kas A., Larbig K., Lim R.M.,  
 RA Smith K.A., Spencer D.H., Wong G.K.-S., Wu Z., Paulsen I.T.,  
 RA Reizer J., Seier M.H., Hancock R.E.W., Lory S., Olson M.V.,  
 RT "Complete genome sequence of Pseudomonas aeruginosa PA01, an  
 opportunistic pathogen.";  
 RL Nature 406:959-964(2000).  
 DR EMBL; AE004551; AAG04603.1; -  
 DR InterPro; IPR000583; GATase\_2.  
 DR PROSITE; PS00443; GATASE\_TYPE\_II; 1.  
 KM Hypothetical protein; Complete proteome.  
 SQ SEQUENCE 532 AA; 59278 MW; A08F8E32DAD28425 CRC64;

Query Match 10.8%; Score 76.5; DB 16; Length 532;  
 Best Local Similarity 26.7%; Pred. No. 17;  
 Matches 40; Conservative 18; Mismatches 67; Indels 25; Gaps 6;  
 QY 4 LCRDSDALTA-----LDEETLWEMESHRHIVRC-----ICPSRLTPYLRQ 45

DB 197 LCORRLATAPASEVDSIDEQ-LGETLERCRQTFPCALLVSGVDSNLLGSYLDPOLQR 255

QY 46 AVVLQQLDEEVLSPLRTNSAMRAGHLLDLKTR-GRNGAIAFLSKFKHPNPDVYTLV 104

DB 256 FHLCHGEDESLPFRLOREFELROEAFMPLIRRAVGNFGATMSILMTQ---RLAD 311

QY 105 GLQPDVFNFGSSDPDGLAGTSRNRL 134  
 DB 312 GIGEGCHVCLLGEAD-ELFWGYRRLLEL 340

## RESULT 7

Q9YW93 ID 09YW93 PRELIMINARY; PRT; 887 AA.  
 AC 09YW93;  
 DT 01-MAY-1999 (TRENBLREL. 10, Created)  
 DT 01-MAY-1999 (TRENBLREL. 10, Last sequence update)  
 DT 01-DEC-2001 (TRENBLREL. 19, Last annotation update)  
 DE Glycoprotein B.  
 GN UL27.  
 OS Macropodid herpesvirus type 1 (parma wallaby herpesvirus).  
 OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;  
 OC Unclassified Herpesviridae.  
 OX NCBI\_TaxID=83441;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA MEDLINE=99171712; PubMed=10073704;  
 RA Mahony T.J., Smith G.A., Thomson D.M.;  
 RT "Macropodid herpesviruses 1 and 2 occupy unexpected molecular  
 phylogenetic positions within the Alphaherpesvirinae.";  
 RL J. Gen. Virol. 80:433-436(1999).  
 DR EMBL; AF061754; AAD11960.1; -  
 DR InterPro; IPR000234; Glycoprot\_B.  
 DR Pfam; PF00606; Glycoprotein B; 1.  
 DR Prodom; PD000693; Glycoprot\_B; 1.  
 SQ SEQUENCE 887 AA; 99739 MW; 0EE4FB4E098F128 CRC64;

Query Match 10.5%; Score 75.5; DB 12; Length 887;  
 Best Local Similarity 28.1%; Pred. No. 41;  
 Matches 25; Conservative 18; Mismatches 37; Indels 9; Gaps 2;

QY 49 LCQDDEEVLSPLRTNSAMRAGHLLDLKTRGKNGAIAFLSKFKHPNPDVYTLVGIQ 108  
 DB 658 ITMEDHEFVPLEVYTRHRIKXSGILDYEVQRNQ-----LHALRFHID-----TIRP 708  
 QY 109 DVDPSNFGSSDPDGLAGTSRNRLLYT 137  
 DB 709 DRYSAIFSGLYSFFDGLABIGRAVAVVT 737

## RESULT 8

Q9FSW0 ID 09FSW0 PRELIMINARY; PRT; 394 AA.  
 AC 09FSW0;  
 DT 01-MAR-2001 (TRENBLREL. 16, Created)  
 DT 01-OCT-2001 (TRENBLREL. 18, Last sequence update)  
 DT 01-JUN-2002 (TRENBLREL. 21, Last annotation update)  
 DE Para-hydroxy benzoate hydroxylase (EC 1.14.13.2).  
 GN POB4.  
 OS Pseudomonas indica.  
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;  
 OC Pseudomonadaceae; Pseudomonas.  
 OX NCBI\_TaxID=137658;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA STRAIN=IMT40;  
 RA Pandey K.K., Mathur A.S., Banasi P.K., Chakrabarti T.;  
 RT "Characterization of genes involved in hydrocarbon utilization in  
 Pseudomonas sp. IMT40.";  
 RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AF302797; AAG17455.2; -  
 DR HSSP; P00438; IPSE.  
 DR InterPro; IPR000733; Flav\_monooxygenase.



DR InterPro: IPR002938; MoxY\_FAD\_binding.  
 DR InterPro: IPR003042; Rng\_moxYgenase.  
 DR Pfam: PF01494; FAD\_binding\_3; 1.  
 DR Pfam: PF01360; Monooxygenase; 1.  
 DR PRINTS: PR00420; RINGMOXGNASE.  
 KM Oxidoreductase.  
 SQ SEQUENCE 394 AA; 44289 MW; F5E9354AE87A27A CRC64;

Query Match 10.3%; Score 74.5; DB 2; Length 394;  
 Best Local Similarity 24.8%; Pred. No. 19;  
 Matches 32; Conservative 17; Mismatches 49; Indels 31; Gaps 4;

QY 29 RIVRCICPSRLTYLFOAKYLQDDEEVLHS-----PRLTNSAMRAGHLL----- 74  
 DB 42 RIRAGVLEQGMVLLREAGVGQMDREGVHDGFELAFDGRLEIRIDRLTGTGKTMVYG 101  
 QY 75 -----DLKTRGKNGAIALLES-----LKFHNDVYTLVTGLQPDVDFNSFGSSSD 121  
 DB 102 QTEVTDLMERASGAPCFYEASEVELHEHLKGENPHVFLHGGQMRIDCOQIAG----- 157  
 QY 122 FDGAGTSR 130  
 DB 158 CDGFGVSR 166

RESULT 9  
 066677 PRELIMINARY; PRT; 210 AA.

AC 066677; (TREMBlrel. 01, Created)  
 DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)  
 DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)  
 DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)  
 DE ORF B10  
 OS Equine herpesvirus type 2 (strain 86/87) (EHV-2).  
 OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;  
 OC Gammaherpesvirinae.  
 CX NCBI\_TaxID=82831;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=9302501; PubMed=7783207;  
 RA Telford E.A., Watson M.S., Aird H.C., Perry J., Davison A.J.;  
 RT "The DNA sequence of equine herpesvirus 2.";  
 RL J. Mol. Biol. 249:520-526(1995).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RA Telford E.A.R.;  
 RL Submitted (FEB-1995) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; U20824; AAC13865.1;  
 DR InterPro: IPR001315; CARD.  
 DR Pfam; PF00619; CARD; 1.  
 DR PROSITE; PS50209; CARD; 1.  
 SQ SEQUENCE 210 AA; 22943 MW; 136D639D8BC2D40 CRC64;

Query Match 10.3%; Score 74; DB 12; Length 210;  
 Best Local Similarity 28.0%; Pred. No. 9.6;  
 Matches 30; Conservative 14; Mismatches 27; Indels 36; Gaps 6;

QY 2 GELCRDSALTALDEETMEM-----MESHRRIVRCICPSRLTPYLRQA 46  
 DB 10 GDPC-----VLTREBIDWVERLCBELRYLVLSHKSHK-----LDHLRAK 52  
 QY 47 KYLCQDDEEVLHSPRLTNSAMRAGHLLDLKTRGKNGAIAFLSLK 93  
 DB 53 KILSRDAEYV--SSRAT--SRSRAGLLVDMCQDHP--GFQCLKESCK 95

RESULT 10  
 09YUN5 PRELIMINARY; PRT; 311 AA.  
 AC 09YUN5; (T-EMBlrel. 10, Created)  
 DT 01-MAY-1999 (T-EMBlrel. 10, Last sequence update)  
 DT 01-MAY-1999 (T-EMBlrel. 10, Last sequence update)  
 DT 01-MAR-2003 (T-EMBlrel. 23, Last annotation update)

DE Hypothetical 32.6 kDa protein.  
 GN E10 OR CLAP.  
 OS Equine herpesvirus 2.  
 OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;  
 OC Gammaherpesvirinae.  
 CX NCBI\_TaxID=133899;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=99142601; PubMed=9989495;  
 RA Willis T.G., Jadayel D.M., Du M.Q., Peng H., Perry A.R.,  
 RA Abdul-Rauf M., Price H., Karren L., Majekodunni O., Wlodarska I.,  
 RA Pan L., Crook T., Hamoudi R., Isaacson P., Dyer M.J.;  
 RT "Bcl10 is involved in multiple tumor types";  
 RL Cell 96:35-45(1999).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RA Thome M., Martinon F., Hofmann K., Rubio V., Steiner V., Schneider P.,  
 RA Matman C., Tschopp J.;  
 RT "Equine herpesvirus-2 B10, but not its cellular homologue, activates  
 RT NF-kB transcription factor and c-Jun N-terminal kinase";  
 RL J. Biol. Chem. 0:0-0(1999).  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=99292766; PubMed=10364242;  
 RA Srinivasula S.M., Ahmad M., Lin J.-H., Royet J.-L.,  
 RA Fernandes-Alnemri T., Trichlia P.N., Alnemri E.S.;  
 RT "CLAP, a novel caspase recruitment domain-containing protein in the  
 RT tumor necrosis factor receptor pathway, regulates NF-kappaB activation  
 RT and apoptosis";  
 RL J. Biol. Chem. 274:17946-17954(1999).  
 DR EMBL; AJ006410; CA07016.1;  
 DR EMBL; AF100340; AAD16430.1;  
 DR EMBL; AF134394; AAD39146.1;  
 DR InterPro: IPR001315; CARD.  
 DR Pfam; PF00619; CARD; 1.  
 DR PROSITE; PS50209; CARD; 1.  
 KM Hypothetical protein.  
 SQ SEQUENCE 311 AA; 32602 MW; 04EC68C7353D723B CRC64;

Query Match 10.3%; Score 74; DB 12; Length 311;  
 Best Local Similarity 28.0%; Pred. No. 16;  
 Matches 30; Conservative 14; Mismatches 27; Indels 36; Gaps 6;

QY 2 GELCRDSALTALDEETMEM-----MESHRRIVRCICPSRLTPYLRQA 46  
 DB 10 GDPC-----VLTREBIDWVERLCBELRYLVLSHKSHK-----LDHLRAK 52  
 QY 47 KYLCQDDEEVLHSPRLTNSAMRAGHLLDLKTRGKNGAIAFLSLK 93  
 DB 53 KILSRDAEYV--SSRAT--SRSRAGLLVDMCQDHP--GFQCLKESCK 95

RESULT 11  
 09NAI9 PRELIMINARY; PRT; 315 AA.  
 AC 09NAI9; (TREMBlrel. 15, Created)  
 DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)  
 DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)  
 DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)  
 DE Y38E10A.22 protein.  
 GN Y38E10A.22.  
 OS Caenorhabditis elegans.  
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae;  
 OC Rhabditidae; Rhabditinae; Caenorhabditis.  
 CX NCBI\_TaxID=6239;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Wallis J.M.;  
 RL Submitted (MAY-1999) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=99069613; PubMed=9851916;

RA none;  
 RT "Genome sequence of the nematode *C.elegans*: A platform for  
 RL investigating biology.";  
 RL Science 282:2012-2018 (1998).  
 DR EMBL: AL110484; CAB54414.1; --  
 DR WormPeP: Y38E10A.22; CER1603.  
 SQ SEQUENCE 315 AA; 36286 MW; 6f69223DFED00831 CRC64;  
 Query Match 10.2%; Score 73.5; DB 5; Length 315;  
 Best Local Similarity 27.0%; Pred. No. 18;  
 Matches 31; Conservative 15; Mismatches 42; Indels 27; Gaps 5;  
 QY 23 MESHRHRIYRCICPSRL--TPYLROAKVLCQDDEEVLHSPRLTNSAMDA-----GHLD 75  
 DB 18 LKNNPEPQALADSLFTAYKKLKRQAL-----SPRLSLALSLSIYASHLPA 70  
 QY 76 LKTRGKNGALA--FLBSLKFNPDV-----YTLVTGLQPDVDFSNFSG 117  
 DB 71 YLKPAKNGLTATEKLIQSTKDKHNAVYAGRYVSLVTLTVKKNPVPFPADLQ 125  
 RESULT 12  
 Q8R6B3 PRELIMINARY; PRT; 382 AA.  
 AC Q8R6B3;  
 DT 01-JUN-2002 (TREMBlrel. 21, Created)  
 DT 01-JUN-2002 (TREMBlrel. 21, Last sequence update)  
 DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)  
 DE (R)-2-hydroxyglutaryl-CoA dehydratase beta-subunit (EC 4.2.1.-).  
 GN FN0208.  
 OS Fusobacterium nucleatum (subsp. nucleatum).  
 OC Bacteria; Fusobacteriales; Fusobacteriaceae;  
 OC Fusobacterium.  
 NCBI\_Taxid=76856;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=ATCC 25586;  
 RX MEDLINE=2186594; PubMed=11889109;  
 RA Kapural V., Anderson I., Ivanova N., Reznik G., Los T., Lykidis A.,  
 RA Bhattacharya A., Bartman A., Gardner W., Greekkin G., Zhu L.,  
 RA Vasileva O., Chu L., Kogan Y., Chaga O., Goldsman E., Bernal A.,  
 RA Larsen N., D'Souza M., Walunas T., Pusch G., Haseikorn R.,  
 RA Fongstein M., Kyriades N., Overbeek R.;  
 RT "Genome sequence and analysis of the oral bacterium *Fusobacterium*  
 RT nucleatum strain ATCC 25586.";  
 RL J. Bacteriol. 184:2005-2016 (2002).  
 DR EMBL: AE010533; AAL94414.1; --  
 KW Lyase; Complete proteome.  
 SQ SEQUENCE 382 AA; 43923 MW; 93038D4296AE34CD CRC64;  
 Query Match 10.2%; Score 73.5; DB 16; Length 382;  
 Best Local Similarity 26.9%; Pred. No. 23;  
 Matches 28; Conservative 15; Mismatches 44; Indels 17; Gaps 4;  
 QY 41 PYLROAKVLCQDDEEVLHSPRLTNSAMRAGHLLDLTKRNGKNGALF----- 89  
 DB 281 PMLALADQFARNDPRLVDPDIK---RPKYVLDLVENNADGGLFMNPNDEEWEY 337  
 QY 90 ESLK--FNNPDVYTLVTGL-QPDVDFSNFSGSSDFDIAGTSR 130  
 DB 338 PSLKQAFDEAKVPLIMGYDQMDVFGQVKQTGLTFENELVQLSR 381  
 RESULT 13  
 Q97EL5 PRELIMINARY; PRT; 265 AA.  
 AC Q97EL5;  
 DT 01-OCT-2001 (TREMBlrel. 18, Created)  
 DT 01-OCT-2001 (TREMBlrel. 18, Last sequence update)  
 DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)  
 DE Hydroxyethylthiazole kinase Thim/Thik (F81).  
 GN CAC3095.  
 OS Clostridium acetobutylicum.

OC Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;  
 OC Clostridium.  
 NCBI\_Taxid=1488;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=ATCC 824 / DSM 792 / VKM B-1787;  
 RX MEDLINE=21359325; PubMed=11466266;  
 RA Noelling U., Breston G., Omelchenko M.V., Makarova K.S., Zeng Q.,  
 RA Gibson R., Lee H.W., Dubois J., Qiu D., Hitti J., Wolf Y.I.,  
 RA Tatusov R.L., Sabathe F., Doucette-Stamm L., Soucaille P., Daly M.J.,  
 RA Bennett G.N., Koonin E.V., Smith D.R.;  
 RT "Genome sequence and comparative analysis of the solvent-producing  
 RT bacterium *Clostridium acetobutylicum*.";  
 RL J. Bacteriol. 183:4823-4838 (2001).  
 DR EMBL: AE007806; AAK81035.1; --  
 DR InterPro: IPR004399; HMP-P\_kinase.  
 DR TrGFams; TIGR00097; HMP-P\_kinase; 1.  
 KW Kinase; Complete proteome.  
 SQ SEQUENCE 265 AA; 28539 MW; 327DA3CEABED06EB CRC64;  
 Query Match 10.1%; Score 73; DB 16; Length 265;  
 Best Local Similarity 23.8%; Pred. No. 16;  
 Matches 20; Conservative 23; Mismatches 29; Indels 12; Gaps 4;  
 QY 29 RIVRCICP--SRTPYLROAKVLCQDDEEVLHSPRLTNS-----AMRAGHLD-L 76  
 DB 125 KTLTCLPLDCLTITPPIPEAVLGGFKIKSQMVAARKISKNVIGILIKGHHVDA 184  
 QY 77 LKTRGKNGALAFBSLKFNPDV 100  
 DB 185 TDLFYKNGYLEWFLKXINDPNTH 208  
 RESULT 14  
 Q8CV98 PRELIMINARY; PRT; 447 AA.  
 AC Q8CV98;  
 DT 01-MAR-2003 (TREMBlrel. 23, Created)  
 DT 01-MAR-2003 (TREMBlrel. 23, Last sequence update)  
 DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)  
 DE NADH oxidase.  
 GN OB0859.  
 OS Oceanobacillus thelyensis.  
 OC Bacteria; Firmicutes; Bacillales; Oceanobacillus.  
 NCBI\_Taxid=182710;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=HTE831 / DSM 14371 / JCM 11309;  
 RX MEDLINE=2220767; PubMed=12235376;  
 RA Takami H., Takaki Y., Uchiyama I.;  
 RT "Genome sequence of *Oceanobacillus thelyensis* isolated from the Ineya  
 RT Ridge and its unexpected adaptive capabilities to extreme  
 RT environments.";  
 RL Nucleic Acids Res. 30:3927-3935 (2002).  
 DR EMBL: AP004596; BAC12815.1; --  
 KW Complete proteome.  
 SQ SEQUENCE 447 AA; 48784 MW; 0388A6FEBB464773 CRC64;  
 Query Match 10.1%; Score 73; DB 16; Length 447;  
 Best Local Similarity 26.8%; Pred. No. 32;  
 Matches 26; Conservative 15; Mismatches 50; Indels 6; Gaps 1;  
 QY 21 EMMESHRIYRCICPSRLTPYLROAKVLCQDDEEVLHSPRLTNSAMRAGHLLDLTKR 80  
 DB 163 EMASLCHRGFSCSLVD-----RSEHYLKRIKEMAHIDELQKGIALVYNDGLKSF 216  
 QY 81 GKNGALAFBSLKFNPDVYTLVTGLQPDVDFSNFSG 117  
 DB 217 SDNGTLLSSDKTIQADMTIMAGIKPNTETLAIDAG 253  
 RESULT 15  
 Q76407

```

ID 076407 PRELIMINARY; PRT; 295 AA.
AC 076407;
DT 01-NOV-1998 (TRENBLREL. 08, Created)
DT 01-OCT-2002 (TRENBLREL. 22, Last sequence update)
DT 01-MAR-2003 (TRENBLREL. 23, Last annotation update)
DE Hypothetical protein T10B5.3.
GN T10B5.3.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditioidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RX MEDLINE=99069613; PubMed=9851916;
RA Waterston R.;
RT "Genome sequence of the nematode C. elegans: a platform for
   investigating biology. The C. elegans Sequencing Consortium.";
RL Science 282:2012-2018(1998).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RA Geisel C.; Bradshaw H.; O'Brien D.;
RT "The sequence of C. elegans cosmid T10B5.";
RL Submitted (JUN-1998) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RA Waterston R.;
RT Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF067947; AAK68397.2; -.
DR Wormpep; T10B5.3; CE30832.
KM Hypothetical protein.
SQ SEQUENCE 295 AA; 32436 MW; 637C20FC0055AA02 CRC64;

```

```

Query March 10.1%; Score 72.5; DB 5; Length 295;
Best Local Similarity 24.3%; Pred. No. 21;
Matches 35; Conservative 20; Mismatches 48; Indels 41; Gaps 6;

```

```

QY 8 DSALTALDEFTLWEMESHRRIVRCIPSRLLTPYLR-----QAKYLCQLDREEV 57
Db 100 DSVIRNINSEII-----HATLRDLQPMQVPLPKITSRUKTRNADIRPTIRWQI 151
QY 58 ---LHSPRLTN-----SAMRAGHLIDLKTRGKNGAIAFLESLEKHPDVY 100
Db 152 AFSIHMEFYLSSLPTLKEKIGGLIGMLRSRVGHHRDLALHGKISTIDLIKRRRTNN----- 207
QY 101 TLVTGLQPDVDFSN-FSGESSDFD 123
Db 208 -VVIQQPLVFNNDLDSSEDFD 230

```

```

Search completed: February 18, 2004, 04:46:49
Job time : 91 secs

```

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using SW model

Run on: February 18, 2004, 04:43:07; Search time 28 Seconds  
(without alignments)  
210.043 Million cell updates/sec

Title: US-10-032-159a-16

Perfect score: 720  
1 MGBLGRDASALTALDEETLW.....SDPDGLAGTSRNRLVTPX 139

Scoring table: BLOSUM62  
Gapop 10.0, Gapext 0.5

Searched: 328717 seqs, 42310858 residues

Total number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database: 1: /cgn2\_6/ptodata/2/iaa/5A.COMB.pep:\*  
2: /cgn2\_6/ptodata/2/iaa/5B.COMB.pep:\*  
3: /cgn2\_6/ptodata/2/iaa/6A.COMB.pep:\*  
4: /cgn2\_6/ptodata/2/iaa/6B.COMB.pep:\*  
5: /cgn2\_6/ptodata/2/iaa/PTCUS.COMB.pep:\*  
6: /cgn2\_6/ptodata/2/iaa/backfile1.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	83	11.5	233	4	US-09-069-023-36
2	78	10.8	233	4	US-09-069-023-38
3	76.5	10.6	621	4	US-09-252-991A-19231
4	76	10.6	326	6	5268278-3
5	74	10.3	333	4	US-09-328-352-7785
6	73.5	10.2	1113	4	US-09-618-425-9
7	72.5	10.1	706	4	US-09-134-001C-4908
8	70.5	9.8	614	4	US-09-252-991A-27778
9	68.5	9.7	1452	2	US-08-652-971-4
10	68.5	9.7	1452	2	US-08-991-258A-4
11	69.5	9.7	1452	2	US-08-769-399-4
12	69.5	9.7	1452	3	US-08-991-953A-4
13	69	9.6	203	4	US-09-198-452A-196
14	69	9.6	603	3	US-09-198-122-2
15	69	9.6	967	3	US-09-139-802-201
16	69	9.6	967	4	US-09-659-786-201
17	68	9.4	164	4	US-09-252-991A-30496
18	68	9.4	313	4	US-09-252-991A-20727
19	67.5	9.4	409	2	US-09-031-485-15
20	67.5	9.4	409	2	US-08-847-429A-15
21	67.5	9.4	409	3	US-09-065-474-15
22	67.5	9.4	409	3	US-09-557-034-15
23	67.5	9.4	422	3	US-09-065-474-142
24	67.5	9.4	422	4	US-09-557-034-142
25	67.5	9.4	497	1	US-08-252-492-2
26	67.5	9.4	497	2	US-08-727-126-2
27	67.5	9.4	497	3	US-08-942-761-2

28	67.5	9.4	1420	2	US-08-540-804-14	Sequence 14, Appl
29	67.5	9.4	1420	2	US-08-218-265-14	Sequence 14, Appl
30	67.5	9.4	1420	3	US-08-521-872-14	Sequence 14, Appl
31	67.5	9.4	1420	3	US-08-590-399-14	Sequence 14, Appl
32	67.5	9.4	1745	2	US-09-031-485-33	Sequence 33, Appl
33	67.5	9.4	1745	2	US-08-847-429A-33	Sequence 33, Appl
34	67.5	9.4	1745	3	US-09-065-474-33	Sequence 33, Appl
35	67.5	9.4	1745	4	US-09-557-034-33	Sequence 33, Appl
36	67	9.3	369	4	US-09-519-232-74	Sequence 74, Appl
37	67	9.3	369	4	US-09-328-352-6146	Sequence 6146, Ap
38	66.5	9.2	345	4	US-09-194-905-12	Sequence 12, Appl
39	66.5	9.2	377	4	US-09-252-991A-32712	Sequence 32712, A
40	66	9.2	355	4	US-09-328-352-4560	Sequence 4560, Ap
41	66	9.2	426	4	US-09-328-352-6412	Sequence 6412, Ap
42	66	9.2	667	4	US-09-342-647-28	Sequence 28, Appl
43	66	9.2	669	4	US-09-342-647-18	Sequence 18, Appl
44	66	9.2	700	4	US-09-621-816B-2	Sequence 2, Appl
45	66	9.2	2532	4	US-09-215-694-10	Sequence 10, Appl

## ALIGNMENTS

RESULT 1  
US-09-069-023-36  
; Sequence 36, Application US/09069023A  
; Patent No. 6348573  
; GENERAL INFORMATION:  
; APPLICANT: Nunez, Gabriel  
; APPLICANT: Inohara, Naohiro  
; APPLICANT: Koseki, Takeyoshi  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR IDENTIFYING APOPTOSIS  
; FILE REFERENCE: UM-03333  
; CURRENT APPLICATION NUMBER: US/09/069,023A  
; CURRENT FILING DATE: 1998-04-27  
; NUMBER OF SEQ ID NOS: 38  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 36  
; LENGTH: 233  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-069-023-36

Query Match 11.5%; Score 83; DB 4; Length 233;

Best Local Similarity 28.5%; Pred. No. 0.02; Matches 39; Conservative 20; Mismatches 46; Indels 32; Gaps 6;

US-09-069-023-36  
; Sequence 36, Application US/09069023A  
; Patent No. 6348573  
; GENERAL INFORMATION:  
; APPLICANT: Nunez, Gabriel  
; APPLICANT: Inohara, Naohiro  
; APPLICANT: Koseki, Takeyoshi  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR IDENTIFYING APOPTOSIS  
; FILE REFERENCE: UM-03333  
; CURRENT APPLICATION NUMBER: US/09/069,023A

CURRENT FILING DATE: 1998-04-27  
 NUMBER OF SEQ ID NOS: 38  
 SOFTWARE: Patentin Ver. 2.0  
 SEQ ID NO 38  
 LENGTH: 233  
 TYPE: PRT  
 ORGANISM: Mus musculus  
 US-09-069-023-38

Query Match 10.8%; Score 78; DB 4; Length 233;  
 Best Local Similarity 35.3%; Pred. No. 0.09;  
 Matches 30; Conservative 14; Mismatches 33; Indels 8; Gaps 4;

QY 13 ALDEETIEM-...MESHRRIYRCIPSRITPYLRQAKVLCQDEEYVLSRLTNSAM 68  
 DB 7 SLTEEDLEVEKKALLENIRVYLCEKIIAEHFHDLRAKKILSREDTEI--SCR-TSSRX 63  
 QY 69 RAGHLDLKTRGKGAIAFLSLK 93  
 DB 64 RAGKLDVYLQ-ENPRGLDTEVSR 87

RESULT 3  
 US-09-252-991A-19231  
 Sequence 19231, Application US/09252991A  
 Patent No. 6551795  
 GENERAL INFORMATION:  
 APPLICANT: Marc J. Rubenfield et al.  
 TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS  
 FILE REFERENCE: 107196.136  
 CURRENT APPLICATION NUMBER: US/09/252,991A  
 PRIOR FILING DATE: 1998-02-18  
 PRIOR APPLICATION NUMBER: US 60/074,788  
 PRIOR FILING DATE: 1998-02-18  
 PRIOR APPLICATION NUMBER: US 60/094,190  
 PRIOR FILING DATE: 1998-07-27  
 NUMBER OF SEQ ID NOS: 33142  
 SEQ ID NO 19231  
 LENGTH: 621  
 TYPE: PRT  
 ORGANISM: Pseudomonas aeruginosa  
 US-09-252-991A-19231

Query Match 10.6%; Score 76.5; DB 4; Length 621;  
 Best Local Similarity 26.7%; Pred. No. 0.63;  
 Matches 40; Conservative 18; Mismatches 67; Indels 25; Gaps 6;

QY 4 LCRASALTA-----LDEETIMMESHRRIYRC-----ICPSRLTPYLQ 45  
 DB 286 LQORRLATAPASVDSDEQ-LSFTLRCKQTPFCALVSGVDVSLDSYLDPOLQR 344  
 QY 46 AKVLCQDEEYVLSRLTNSAMRAGHLDLKTR-GRGALAFLESKFNPDVYTLVT 104  
 DB 345 FHLCEGEBESILPRRLQRFELRQAFMPLBRVGNFGATRMSILMTQ----RLAD 400  
 QY 105 GLQPDVDFSNFSGESSDPDGLAGTSRNLRL 134  
 DB 401 GIGEGYHCVLIGEGAD-ELFWGYPRHLEL 429

RESULT 4  
 5268278-3  
 Patent No. 5268278  
 APPLICANT: CANOSI, UMBERTO, DE FAZIO, GABRIELE, VILIA, STEFANO;  
 DONINI, SILVIA  
 TITLE OF INVENTION: GENETIC EXPRESSION OF SOMATOSTATIN AS  
 HYBRID POLYPEPTIDE  
 NUMBER OF SEQUENCES: 4  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/07/193,202  
 FILING DATE: 09-MAY-1988  
 PRIOR APPLICATION DATA:

APPLICATION NUMBER: 717,444  
 FILING DATE: 29-MAR-1985  
 SEQ ID NO:3;  
 LENGTH: 326  
 5268278-3

Query Match 10.6%; Score 76; DB 6; Length 326;  
 Best Local Similarity 27.0%; Pred. No. 0.27;  
 Matches 37; Conservative 22; Mismatches 58; Indels 20; Gaps 7;

QY 9 SALTAIDEB--TLWEMESHRRIVRCIPSRITPYLRQAKVLCQD-EEYVLSRL 63  
 DB 195 ASLFAPDEEEKRLTARLDELRLQQLTEAP--LPVVSVPKRCQCSDEEFGVRL 251  
 QY 64 TNSAMRAGHLDLKTRGN---GALAFLESKFNPDVYTLVTGLQPDVDFSNF--SG 117  
 DB 252 LQKAIKAGEIFQVVSRRFSLPCPSPLAAYVLLKSDSPYWF--FMQDNDFTLFGASP 308  
 QY 118 ESSDPDGLAGTSRNLRL 134  
 DB 309 ESSKD--ATSRQIEI 322

RESULT 5  
 US-09-328-352-7785  
 Sequence 7785, Application US/09328352  
 Patent No. 6562958  
 GENERAL INFORMATION:  
 APPLICANT: Gary L. Breton et al.  
 TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER  
 FILE REFERENCE: GTC99-03PA  
 CURRENT APPLICATION NUMBER: US/09/328,352  
 PRIOR FILING DATE: 1999-06-04  
 NUMBER OF SEQ ID NOS: 8252  
 SEQ ID NO 7785  
 LENGTH: 333  
 TYPE: PRT  
 ORGANISM: Acinetobacter baumannii  
 US-09-328-352-7785

Query Match 10.3%; Score 74; DB 4; Length 333;  
 Best Local Similarity 27.9%; Pred. No. 0.51;  
 Matches 39; Conservative 14; Mismatches 41; Indels 46; Gaps 8;

QY 11 LTAIDETIMMESHRRIYRCI-----CPSRLTPY-----LRQAKVLCQDEE 56  
 DB 135 ITSDVDDLDHGGAAH--FVKCIEIRNRCPDTLIEILVDPDRGLTALSTLSLSPD 191  
 QY 57 VLS-----PRLTNSAMRAG--HLLDLK-----TRGNG-----ALAF 88  
 DB 192 VFNENETVYRL-YKAMRPSGDYGSINLLKPKKVCPIKRGGLMVGGEIARVIL 250  
 QY 89 LESLKFNPDVYTLVTGLQ 108  
 DB 251 LNDLKDHVDVLTITIGYQLQ 270

RESULT 6  
 US-09-618-425-9  
 Sequence 9, Application US/09618425  
 Patent No. 6475744  
 GENERAL INFORMATION:  
 APPLICANT: Reppert, Steven M.  
 APPLICANT: Weaver, David R.  
 APPLICANT: Zyika, Mark  
 APPLICANT: Jin, Xiaowei  
 APPLICANT: Kume, Kazuhiko  
 APPLICANT: Siram, Satyanarayanan  
 APPLICANT: Shearman, Lauren  
 TITLE OF INVENTION: METHODS FOR IDENTIFYING COMPOUNDS WHICH  
 TITLE OF INVENTION: MODULATE CIRCADIAN RHYTHM  
 FILE REFERENCE: 00786-428001

; CURRENT APPLICATION NUMBER: US/09/618,425  
 ; CURRENT FILING DATE: 2000-07-18  
 ; PRIOR APPLICATION NUMBER: 60/203,005  
 ; PRIOR FILING DATE: 2000-05-10  
 ; PRIOR APPLICATION NUMBER: 60/145,363  
 ; PRIOR FILING DATE: 1999-07-22  
 ; NUMBER OF SEQ ID NOS: 13  
 ; SOFTWARE: FastSeq for Windows Version 4.0  
 ; SEQ ID NO 9  
 ; LENGTH: 1113  
 ; TYPE: PRT  
 ; ORGANISM: Mus musculus  
 ; US-09-618-425-9

Query Match 10.2%; Score 73.5; DB 4; Length 1113;  
 Best Local Similarity 25.2%; Pred. No. 3.8;  
 Matches 26; Conservative 19; Mismatches 33; Indels 25; Gaps 4;

QY 6 RDSALTLADEETLWEMESHRRIVRCICPSRLTPYKAKVLCQ-----LDEEVL 58  
 DB 1003 QDEELPFAABESIRWME---RTPECVMTYQVPERGREEVKODLEKLSMEQOQPL 1058  
 QY 59 HSPRL-----TNSAMRAGHLIDL--KTRGNKGAIA 87  
 DB 1059 FSPQREELAKVRSWISHHTAPQESHTQSCVACEDRGSVGDTA 1101

RESULT 7  
 ; US-09-134-001C-4908  
 ; Sequence 4908, Application US/09134001C  
 ; Patent No. 6380370  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Lynn Doucette-Stamm et al  
 ; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS  
 ; TITLE OF INVENTION: EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS  
 ; FILE REFERENCE: GTC-007  
 ; CURRENT APPLICATION NUMBER: US/09/134,001C  
 ; CURRENT FILING DATE: 1998-08-13  
 ; PRIOR APPLICATION NUMBER: US 60/064,964  
 ; PRIOR FILING DATE: 1997-11-08  
 ; PRIOR APPLICATION NUMBER: US 60/055,779  
 ; PRIOR FILING DATE: 1997-08-14  
 ; NUMBER OF SEQ ID NOS: 5674  
 ; SEQ ID NO 4908  
 ; LENGTH: 706  
 ; TYPE: PRT  
 ; ORGANISM: Staphylococcus epidermidis  
 ; US-09-134-001C-4908

Query Match 10.1%; Score 72.5; DB 4; Length 706;  
 Best Local Similarity 25.4%; Pred. No. 2.5;  
 Matches 36; Conservative 18; Mismatches 47; Indels 41; Gaps 6;

QY 6 RDSALTLADEETLWEMESH-----RHRIVRCICPSRLTPYKROA 46  
 DB 271 QDEELDLKKEVYGFHDEEDPENETLYKEVYALINDLIEBVRRLIADKXIPDGRKV 330  
 QY 47 KVLQDLDEEYVLSRLTNSAMRAGHLIDLKTRGNKGAIAFLSKFHNPDVYTLVTL 106  
 DB 331 DEIRPL-BSYVGLLFRAGSG-----LFTNGQTALSVTLGALGD---YQLIDGL 377  
 QY 107 QPDVD-----PSNFS-GES 119  
 DB 378 GPEVEKRFMEHYNFPNFSVGET 399

RESULT 8  
 ; US-09-252-991A-27778  
 ; Sequence 27778, Application US/09252991A  
 ; Patent No. 6551795  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Marc J. Rubenfield et al.  
 ; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS

; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS  
 ; FILE REFERENCE: 107196,136  
 ; CURRENT APPLICATION NUMBER: US/09/252,991A  
 ; CURRENT FILING DATE: 1999-02-18  
 ; PRIOR APPLICATION NUMBER: US 60/074,788  
 ; PRIOR FILING DATE: 1998-02-18  
 ; PRIOR APPLICATION NUMBER: US 60/094,190  
 ; PRIOR FILING DATE: 1998-07-27  
 ; NUMBER OF SEQ ID NOS: 33142  
 ; SEQ ID NO 27778  
 ; LENGTH: 614  
 ; TYPE: PRT  
 ; ORGANISM: Pseudomonas aeruginosa  
 ; US-09-252-991A-27778

Query Match 9.8%; Score 70.5; DB 4; Length 614;  
 Best Local Similarity 25.2%; Pred. No. 3.7;  
 Matches 31; Conservative 17; Mismatches 44; Indels 31; Gaps 5;

QY 46 AKVLCQDLDEE---VLHSP-----RLTNSAMRAGHLIDLKTRG--- 81  
 DB 368 AKPLQGLEEETSDALVNPATRTLTFTGKRPVLVELSTGDLRTPIPLGSPGVAV 427  
 QY 82 -KKGATAFLSK-----TH-NDVYTLVYGLQPDVDFSNFSGSSDPGLACTSRNLK 134  
 DB 428 LKNGNVAVDERNTLTTHVDPTALSTKSLAEPLPLNIGKNGKFGIADWPPQORL 487  
 QY 135 LVT 137  
 DB 488 LLS 490

RESULT 9  
 ; US-08-652-971-4  
 ; Sequence 4, Application US/08652971  
 ; Patent No. 5814507  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Cheng, J111  
 ; APPLICANT: Laeky, Laurence A.  
 ; TITLE OF INVENTION: A NOVEL KAPPA/MU-LIKE PROTEIN TYROSINE  
 ; TITLE OF INVENTION: PHOSPHATASE, PTP LAMEDA  
 ; NUMBER OF SEQUENCES: 10  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: Genentech, Inc.  
 ; STREET: 460 Point San Bruno Blvd.  
 ; CITY: South San Francisco  
 ; STATE: California  
 ; COUNTRY: United States  
 ; ZIP: 94080  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Floppy disk  
 ; COMPUTER: IBM PC compatible  
 ; OPERATING SYSTEM: PC-DOS/MS-DOS  
 ; SOFTWARE: Patent Release #1.0, Version #1.30  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/08/652,971  
 ; FILING DATE:  
 ; CLASSIFICATION: 435  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: Dreger, Ginger R.  
 ; REGISTRATION NUMBER: 33,055  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: (415) 225-3216  
 ; TELEFAX: (415) 952-9881  
 ; TELEX: 910 371-7168  
 ; INFORMATION FOR SEQ ID NO: 4:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 1452 amino acids  
 ; TYPE: amino acid  
 ; STRANDEDNESS:  
 ; TOPOLOGY: linear  
 ; MOLECULE TYPE: protein

US-08-652-971-4

Query Match 9.7%; Score 69.5; DB 2; Length 1452;  
Best Local Similarity 26.9%; Pred. No. 19;  
Matches 32; Conservative 14; Mismatches 50; Indels 23; Gaps 6;

QY 16 EETLWEMESH-RHRIYRCICPSRLPYLROAKYLCOUDEEVLHSPRLTNSAMRAGHLL 74  
DB 431 EEVSWDNDNSHPQHTI-----TNLSPTNVSVKLIIMNPKGKESQELT--VQTDDELIP 482  
QY 75 DLKTRKNGAIAFLBS--LKFNPD-----VYTLVTGLQPDVDFSNFSGESS 120  
DB 483 GAVPTESIQGS-AFEKIFLOWREPQTGYVITLYETTYKAVSFPDEIDLSNOSGRVS 540

RESULT 10  
US-08-991-258A-4  
Sequence 4, Application US/08991258A  
Patent No. 5928887

GENERAL INFORMATION:  
APPLICANT: Cheng, Jili  
APPLICANT: Lasky, Laurence A.  
TITLE OF INVENTION: A NOVEL KAPPA/MU-LIKE PROTEIN TYROSINE  
NUMBER OF SEQUENCES: 10  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: FLEHR, HOBBACH, TEST, ALBRITTON & HERBERT, LLP  
STREET: 4 Embarcadero Center, Suite 3400  
CITY: San Francisco  
STATE: California  
COUNTRY: United States  
ZIP: 94111  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent in Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/991,258A  
FILING DATE: 17-DEC-1997  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/652,971  
FILING DATE: 24-MAY-1996  
ATTORNEY/AGENT INFORMATION:  
NAME: Dreger, Walter H.  
REGISTRATION NUMBER: 24,190  
REFERENCE/DOCKET NUMBER: A-63478-3/WHD/MTX  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (415) 781-1989  
TELEFAX: (415) 398-3249  
TELEX:  
INFORMATION FOR SEQ ID NO: 4:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1452 amino acids  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-991-258A-4

Query Match 9.7%; Score 69.5; DB 2; Length 1452;  
Best Local Similarity 26.9%; Pred. No. 19;  
Matches 32; Conservative 14; Mismatches 50; Indels 23; Gaps 6;  
QY 16 EETLWEMESH-RHRIYRCICPSRLPYLROAKYLCOUDEEVLHSPRLTNSAMRAGHLL 74  
DB 431 EEVSWDNDNSHPQHTI-----TNLSPTNVSVKLIIMNPKGKESQELT--VQTDDELIP 482  
QY 75 DLKTRKNGAIAFLBS--LKFNPD-----VYTLVTGLQPDVDFSNFSGESS 120  
DB 483 GAVPTESIQGS-AFEKIFLOWREPQTGYVITLYETTYKAVSFPDEIDLSNOSGRVS 540

RESULT 11  
US-08-769-399-4  
Sequence 4, Application US/08769399  
Patent No. 5976852

GENERAL INFORMATION:  
APPLICANT: Cheng, Jili  
APPLICANT: Lasky, Laurence A.  
TITLE OF INVENTION: A NOVEL KAPPA/MU-LIKE PROTEIN TYROSINE  
NUMBER OF SEQUENCES: 10  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Genentech, Inc.  
STREET: 460 Point San Bruno Blvd.  
CITY: South San Francisco  
STATE: California  
COUNTRY: United States  
ZIP: 94080  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent in Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/769,399  
FILING DATE:  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Dreger, Ginger R.  
REGISTRATION NUMBER: 33,055  
REFERENCE/DOCKET NUMBER: P1033  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (415) 225-3216  
TELEFAX: (415) 952-9881  
TELEX: 910 371-7168  
INFORMATION FOR SEQ ID NO: 4:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1452 amino acids  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-769-399-4

Query Match 9.7%; Score 69.5; DB 2; Length 1452;  
Best Local Similarity 26.9%; Pred. No. 19;  
Matches 32; Conservative 14; Mismatches 50; Indels 23; Gaps 6;

QY 16 EETLWEMESH-RHRIYRCICPSRLPYLROAKYLCOUDEEVLHSPRLTNSAMRAGHLL 74  
DB 431 EEVSWDNDNSHPQHTI-----TNLSPTNVSVKLIIMNPKGKESQELT--VQTDDELIP 482  
QY 75 DLKTRKNGAIAFLBS--LKFNPD-----VYTLVTGLQPDVDFSNFSGESS 120  
DB 483 GAVPTESIQGS-AFEKIFLOWREPQTGYVITLYETTYKAVSFPDEIDLSNOSGRVS 540

RESULT 12  
US-08-991-953A-4  
Sequence 4, Application US/08991953A  
Patent No. 6083748

GENERAL INFORMATION:  
APPLICANT: Cheng, Jili  
APPLICANT: Lasky, Laurence A.  
TITLE OF INVENTION: A NOVEL KAPPA/MU-LIKE PROTEIN TYROSINE  
NUMBER OF SEQUENCES: 10  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: FLEHR, HOBBACH, TEST, ALBRITTON & HERBERT, LLP  
STREET: 4 Embarcadero Center, Suite 3400  
CITY: San Francisco  
STATE: California  
COUNTRY: United States

ZIP: 94111  
COMPUTER READABLE FORM:  
MEDIUM TYPE: floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/991.953A  
FILING DATE: 16-DEC-1997  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/652,971  
FILING DATE: 24-MAY-1996  
ATTORNEY/AGENT INFORMATION:  
NAME: Dreger, Walter H.  
REGISTRATION NUMBER: 24,190  
REFERENCE/DOCKET NUMBER: A-63478-3/WHD/MTK  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (415) 781-1989  
TELEFAX: (415) 398-3249  
TELEX:  
INFORMATION FOR SEQ ID NO: 4:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1452 amino acids  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-991-953A-4

Query Match 9.7%; Score 69.5; DB 3; Length 1452;  
Best Local Similarity 26.9%; Pred. No. 19;  
Matches 32; Conservative 14; Mismatches 50; Indels 23; Gaps 6;

QY 16 ESTLMEMESH-RHRIVRCIPSRITPYLRQAKVLCQDEEEVLHSPRLTNSAARAGHL 74  
DB 431 EVSMDTNSHQHT-----TMSPTYNVSVKILMNPGRKRSGLT--VQTDLP 482

QY 75 DLKTRGNKGAIAFLES--LKFHPND-----YTLVTGLQPDVPSNFGSS 120  
DB 483 GAVPTESIQGS-APEKIFLQWREPTQYGYITLYEITYKAVSFDPEIDLSNGRVS 540

RESULT 13  
US-09-198-452A-196  
; Sequence 196, Application US/09198452A  
; Patent No. 6559294  
; GENERAL INFORMATION:  
; APPLICANT: Griffiths, R.  
; TITLE OF INVENTION: Chlamydia pneumoniae genomic sequence and polypeptides, fragments  
; TITLE OF INVENTION: thereof and uses thereof, in particular for the diagnosis, prevention  
; FILE REFERENCE: 9710-003-999  
; CURRENT APPLICATION NUMBER: US/09/198,452A  
; CURRENT FILING DATE: 1998-11-24  
; NUMBER OF SEQ ID NOS: 6849  
; SEQ ID NO 196  
; LENGTH: 203  
; TYPE: PRT  
; ORGANISM: Chlamydia pneumoniae  
US-09-198-452A-196

Query Match 9.6%; Score 69; DB 4; Length 203;  
Best Local Similarity 21.1%; Pred. No. 1.1; Mismatches 47; Indels 20; Gaps 3;  
Matches 24; Conservative 23;

QY 28 HRIVACIPSRITPYLRQAKVLCQDEEEVLHSPRLTNSA-----WEAGHLDLK 78  
DB 4 HXILDG-----KXKVFQNTVASSETYPSQILHAGSEVADAFNQADCHPAANQILEAKK 59

QY 79 TRGNKGAIAFLESKLFHPNDYTLVTGLQPDVPSNFGSSDPDGLAGTSRNL 132  
DB 60 -----ICLLDVHTNHYSVFTFCVDNYPNLRFYVSSKNMENGJSLNPDNY 106

RESULT 14  
US-09-198-122-2  
; Sequence 2, Application US/09198122  
; Patent No. 6180380  
; GENERAL INFORMATION:  
; APPLICANT: Streibardt, Klaus; Rubsamen-Waigmann, Helga;  
; APPLICANT: Holtzrich, Uwe  
; TITLE OF INVENTION: CLONING OF A MEMBER OF THE SERINE-  
; TITLE OF INVENTION: THRONINE-KINASE FAMILY  
; NUMBER OF SEQUENCES: 7  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: SPRUNG HORN KRAMER & WOODS  
; STREET: 660 White Plains Road  
; CITY: Tarrytown  
; STATE: New York  
; COUNTRY: U.S.A.  
; ZIP: 10591-5144  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette, 3.50 inch, 2.0 MB  
MEDIUM TYPE: storage  
COMPUTER: NEC Powermate SX-20  
OPERATING SYSTEM: DOS  
SOFTWARE: WordPerfect 5.1  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/198,122  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/08/601,014  
FILING DATE: 23-FEB-1996  
APPLICATION NUMBER: PCT/EP94/02863  
FILING DATE: 30-AUG-1994  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: DE 43292177  
FILING DATE: 30-AUG-1993  
ATTORNEY/AGENT INFORMATION:  
NAME: Kurt G. Briscoe  
REGISTRATION NUMBER: 33,141  
REFERENCE/DOCKET NUMBER: Bayer 9516-KGB  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (914) 332-1700  
TELEFAX: (914) 332-1844  
TELEX:  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 603 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: Protein  
US-09-198-122-2

Query Match 9.6%; Score 69; DB 3; Length 603;  
Best Local Similarity 24.5%; Pred. No. 5.6;  
Matches 34; Conservative 15; Mismatches 38; Indels 52; Gaps 6;

QY 3 ELGRDSALTADEEFLTWMESHRIIVRCIPSRITPYLRQAKVLCQDEEEVLHSPR 62  
DB 131 ELGRRS-----LLEPHKRR--KALTEPARAYLRQIVGQ-----YLHRRN 171

QY 63 LTNASARAGHL--DLKTRGNKGAIAFLE-----SLKFPND 98  
DB 172 VHRDKLGNLFNEDLVKIGDFGLATKVEYDGERKTLGCPNPIABVLSKSHSFE 231

QY 99 V-----YTLVTGLQPD 108  
DB 232 VDWVSGICINVTLLVGKPP 250

RESULT 15  
US-09-139-802-201  
; Sequence 201, Application US/09139802



```

; Patent No. 6180084
; GENERAL INFORMATION:
; APPLICANT: Ruoslahti, Erkki
; APPLICANT: Pasqualini, Renata
; TITLE OF INVENTION: NGR Receptor and Methods of Identifying Tumor Homing
; TITLE OF INVENTION: Molecules That Home to Angiogenic Vasculature using
; TITLE OF INVENTION: Same
; FILE REFERENCE: P-IJ 3203
; CURRENT APPLICATION NUMBER: US/09/139, 802
; CURRENT FILING DATE: 1998-08-25
; EARLIER APPLICATION NUMBER: 08/926, 914
; EARLIER FILING DATE: 1997-09-10
; EARLIER APPLICATION NUMBER: 08/710, 067
; EARLIER FILING DATE: 1996-09-10
; NUMBER OF SEQ ID NOS: 226
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 201
; LENGTH: 967
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-139-802-201

```

```

Query Match          9.6%; Score 69; DB 3; Length 967;
Best Local Similarity 25.0%; Pred. No. 12;
Matches 40; Conservative 25; Mismatches 51; Indels 44; Gaps 11;

QY 9 SALFALDEETIMEMESHRRIRVICPS---RLTPYL-----RQAKYL-C- 50
DB 60 ASATTLDDSKAW---NRYRLPNTLKPDSYQVTLRPYLTPNDRLGLYVFKGSSIVRFTCK 114
QY 51 QLDEEEVLAHSPRLTNSAMRAGHLIDLKTRGNKA-IAFLESLEKFNPDVY-----TL 102
DB 115 EATVILIIHSKGL-NYILSGHRVYL---RQVGGSSQEPDIDKTRLEPTEYLIVHLKSL 170
QY 103 VTGLQPDVDFSNFSGSSDFDGLAGTSR-----NRLILV 136
DB 171 VKDSQYEMD-SEFEGELA--DLAGFYRSEYMEGNVRKV 207

```

Search completed: February 18, 2004, 04:48:20  
 Job time : 31 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 18, 2004, 04:46:58 ; Search time 76 Seconds

(without alignments)  
382.949 Million cell updates/sec

Title: US-10-032-159a-16

Sequence: 1 MGELCRDSALTPALDEETLM.....SDPDLGAGTSRNLRLVTPX 139

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 801455 seqs, 209382283 residues

Total number of hits satisfying chosen parameters: 801455

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database :

Published Applications AA:\*

- 1: /cgn2\_6/ptodata/1/pubpaa/US07\_PUBCOMB.pep:\*
- 2: /cgn2\_6/ptodata/1/pubpaa/PCT\_NEW\_PUB.pep:\*
- 3: /cgn2\_6/ptodata/1/pubpaa/US06\_NEW\_PUB.pep:\*
- 4: /cgn2\_6/ptodata/1/pubpaa/US06\_PUBCOMB.pep:\*
- 5: /cgn2\_6/ptodata/1/pubpaa/US07\_NEW\_PUB.pep:\*
- 6: /cgn2\_6/ptodata/1/pubpaa/PCTUS\_PUBCOMB.pep:\*
- 7: /cgn2\_6/ptodata/1/pubpaa/US08\_NEW\_PUB.pep:\*
- 8: /cgn2\_6/ptodata/1/pubpaa/US08\_PUBCOMB.pep:\*
- 9: /cgn2\_6/ptodata/1/pubpaa/US09\_PUBCOMB.pep:\*
- 10: /cgn2\_6/ptodata/1/pubpaa/US09\_PUBCOMB.pep:\*
- 11: /cgn2\_6/ptodata/1/pubpaa/US09\_NEW\_PUB.pep:\*
- 12: /cgn2\_6/ptodata/1/pubpaa/US09\_PUBCOMB.pep:\*
- 13: /cgn2\_6/ptodata/1/pubpaa/US10\_PUBCOMB.pep:\*
- 14: /cgn2\_6/ptodata/1/pubpaa/US10\_PUBCOMB.pep:\*
- 15: /cgn2\_6/ptodata/1/pubpaa/US10C\_PUBCOMB.pep:\*
- 16: /cgn2\_6/ptodata/1/pubpaa/US10C\_NEW\_PUB.pep:\*
- 17: /cgn2\_6/ptodata/1/pubpaa/US60\_NEW\_PUB.pep:\*
- 18: /cgn2\_6/ptodata/1/pubpaa/US60\_PUBCOMB.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	719	99.9	139	US-10-032-159a-16	Sequence 16, Appl
2	714	99.2	138	US-09-767-215-5	Sequence 5, Appl
3	617.5	85.8	418	US-10-104-047-2285	Sequence 2285, Ap
4	617.5	85.8	1004	US-09-767-215-2	Sequence 2, Appl
5	484	67.2	92	US-10-032-159a-18	Sequence 16, Appl
6	270.5	37.6	1147	US-09-798-412-11	Sequence 11, Appl
7	270.5	37.6	1147	US-10-325-917-11	Sequence 8, Appl
8	270.5	37.6	1247	US-10-032-159a-8	Sequence 10, Appl
9	257	35.7	92	US-10-032-159a-10	Sequence 8, Appl
10	246	34.2	1032	US-10-325-917-8	Sequence 2, Appl
11	246	34.2	1032	US-10-325-917-8	Sequence 2, Appl
12	223.5	31.0	536	US-09-798-412-2	Sequence 2, Appl
13	223.5	31.0	536	US-10-032-159a-2	Sequence 20, Appl
14	215.5	29.9	366	US-10-032-159a-20	
15	215.5	29.9	366	US-10-032-159a-20	

16	215.5	29.9	536	11	US-09-798-412-5	Sequence 5, Appl
17	215.5	29.9	536	15	US-10-325-917-5	Sequence 5, Appl
18	158	21.9	72	14	US-10-032-159a-4	Sequence 4, Appl
19	113	15.7	48	12	US-10-029-386-34276	Sequence 34276, A
20	80	11.1	966	12	US-10-099-322-72	Sequence 72, Appl
21	80	11.1	966	12	US-10-044-564-72	Sequence 58, Appl
22	75.5	10.5	685	11	US-09-877-843-58	Sequence 12060, A
23	74.5	10.3	281	15	US-10-156-761-12060	Sequence 12989, A
24	74.5	10.3	329	12	US-10-369-493-12989	Sequence 6948, Ap
25	73.5	10.2	1113	15	US-10-245-175-9	Sequence 4, Appl
26	72.5	10.1	1318	12	US-10-369-493-6948	Sequence 6408, Ap
27	71.5	9.9	889	11	US-09-099-567B-44	Sequence 196, App
28	70.5	9.8	356	12	US-10-369-493-6408	Sequence 4617, Ap
29	69.5	9.7	431	14	US-10-078-929-200	Sequence 186, App
30	69	9.6	203	12	US-10-289-762-196	Sequence 71, Appl
31	69	9.6	395	12	US-10-108-260A-4617	Sequence 14, Appl
32	69	9.6	603	15	US-10-171-311-186	Sequence 75, Appl
33	69	9.6	966	12	US-10-099-322-71	Sequence 14, Appl
34	69	9.6	966	12	US-10-044-564-71	Sequence 74, Appl
35	69	9.6	967	12	US-10-099-322-14	Sequence 75, Appl
36	69	9.6	967	12	US-10-099-322-74	Sequence 32, Appl
37	69	9.6	967	12	US-10-099-322-75	Sequence 201, Appl
38	69	9.6	967	12	US-10-044-564-14	Sequence 2, Appl
39	69	9.6	967	12	US-10-044-564-74	Sequence 73, Appl
40	69	9.6	967	12	US-10-044-564-75	
41	69	9.6	967	15	US-10-205-823-32	
42	69	9.6	967	15	US-10-264-374-201	
43	69	9.6	967	15	US-10-281-904-2	
44	69	9.6	969	10	US-09-981-353-122	
45	69	9.6	974	12	US-10-099-322-73	

#### ALIGNMENTS

RESULT 1  
US-10-032-159a-16  
; Sequence 16, Application US/10032159A  
; Publication No. US20020164703A1  
; GENERAL INFORMATION:  
; APPLICANT: Pawlowski, Krzysztof  
; APPLICANT: Reed, John C.  
; TITLE OF INVENTION: CARD-DOMAIN CONTAINING POLYPEPTIDES,  
; TITLE OF INVENTION: ENCODING NUCLEIC ACIDS, AND METHODS OF USE  
; FILE REFERENCE: P-1J 5100  
; CURRENT APPLICATION NUMBER: US/10/032.159A  
; PRIOR FILING DATE: 2001-12-19  
; PRIOR APPLICATION NUMBER: US 60/257,457  
; PRIOR FILING DATE: 2000-12-21  
; NUMBER OF SEQ ID NOS: 37  
; SOFTWARE: PasteSeq for Windows Version 4.0  
; SEQ ID NO 16  
; LENGTH: 139  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: VARIANT  
; LOCATION: 139  
; OTHER INFORMATION: Xaa = Any Amino Acid  
US-10-032-159a-16  
Query Match 99.9%; Score 719; DB 14; Length 139;  
Best Local Similarity 100.0%; Pred. No. 6.1e-77;  
Matches 138; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 MGELCRDSALTPALDEETLMEMESHRRIVRCICPSRLTPYVROAKVLCOLBEETVHS 60  
DB 1 MGELCRDSALTPALDEETLMEMESHRRIVRCICPSRLTPYVROAKVLCOLBEETVHS 60  
QY 61 PRLTNSMRAGHLIDLKTRGKNGAIAFLESLEKFNBDVTVTVGLQPDVDFSNFSGESS 120  
DB 61 PRLTNSMRAGHLIDLKTRGKNGAIAFLESLEKFNBDVTVTVGLQPDVDFSNFSGESS 120

Qy 121 DFDGAGTSRNLRLVTP 138  
Db 121 DFDGAGTSRNLRLVTP 138

## RESULT 2

US-09-767-215-5  
; Sequence 5, Application US/09767215  
; Patent No. US20020081636A1  
; GENERAL INFORMATION:  
; APPLICANT: Bertin, John  
; TITLE OF INVENTION: NOVEL MOLECULES OF THE CARD-RELATED  
; FILE REFERENCE: 07334-142001  
; CURRENT APPLICATION NUMBER: US/09/767,215  
; PRIOR FILING DATE: 2001-01-22  
; PRIOR APPLICATION NUMBER: 60/181,159  
; PRIOR FILING DATE: 2000-02-09  
; NUMBER OF SEQ ID NOS: 10  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 5  
; LENGTH: 1138  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-767-215-5

Query Match 99.2%; Score 714; DB 9; Length 1138;  
Best Local Similarity 100.0%; Pred. No. 4,5e-75;  
Matches 137; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 GELCRDSALTALDEETLWMMESHRRIRIVRCICPSRLTPYLROAKYLCQDDEEVVHSP 61  
Db 1 GELCRDSALTALDEETLWMMESHRRIRIVRCICPSRLTPYLROAKYLCQDDEEVVHSP 60  
Qy 62 RLTSNAPRAGHLLDLKTRGNKGAIAFLSEIKFHNPDVYTLVTGLQPDVDFSNFSGSSD 121  
Db 61 RLTSNAPRAGHLLDLKTRGNKGAIAFLSEIKFHNPDVYTLVTGLQPDVDFSNFSGSSD 120  
Qy 122 FDGLAGTSRNLRLVTP 138  
Db 121 FDGLAGTSRNLRLVTP 137

## RESULT 3

US-10-104-047-2285  
; Sequence 2285, Application US/10104047  
; Publication No. US20030236392A1  
; GENERAL INFORMATION:  
; APPLICANT: HELIX RESEARCH INSTITUTE  
; TITLE OF INVENTION: NO. US20030236392A1 full length cDNA  
; FILE REFERENCE: H1-A0105  
; CURRENT APPLICATION NUMBER: US/10/104,047  
; PRIOR FILING DATE: 2002-03-25  
; PRIOR APPLICATION NUMBER:  
; PRIOR FILING DATE:  
; NUMBER OF SEQ ID NOS: 4096  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 2285  
; LENGTH: 418  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-104-047-2285

Query Match 85.8%; Score 617.5; DB 12; Length 418;  
Best Local Similarity 90.4%; Pred. No. 2,9e-64;  
Matches 123; Conservative 4; Mismatches 6; Indels 3; Gaps 2;

Qy 1 MGELCRDSALTALDEETLWMMESHRRIRIVRCICPSRLTPYLROAKYLCQDDEEVVHSP 60  
Db 1 MGELCRDSALTALDEETLWMMESHRRIRIVRCICPSRLTPYLROAKYLCQDDEEVVHSP 60  
Qy 61 RLTSNAPRAGHLLDLKTRGNKGAIAFLSEIKFHNPDVYTLVTGLQPDVDFSNFSG--E 118

Db 61 RLTSNAPRAGHLLDLKTRGNKGAIAFLSEIKFHNPDVYTLVTGLQPDVDFSNFSGME 120  
Qy 119 SSDF-DGLAGTSRNLRL 133  
Db 121 TSKLTCTCAGAGTSLQ 136

## RESULT 4

US-09-767-215-2  
; Sequence 2, Application US/09767215  
; Patent No. US20020081636A1  
; GENERAL INFORMATION:  
; APPLICANT: Bertin, John  
; TITLE OF INVENTION: NOVEL MOLECULES OF THE CARD-RELATED  
; FILE REFERENCE: 07334-142001  
; CURRENT APPLICATION NUMBER: US/09/767,215  
; PRIOR FILING DATE: 2001-01-22  
; PRIOR APPLICATION NUMBER: 60/181,159  
; PRIOR FILING DATE: 2000-02-09  
; NUMBER OF SEQ ID NOS: 10  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 2  
; LENGTH: 1004  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-767-215-2

Query Match 85.8%; Score 617.5; DB 9; Length 1004;  
Best Local Similarity 90.4%; Pred. No. 9,9e-64;  
Matches 123; Conservative 4; Mismatches 6; Indels 3; Gaps 2;

Qy 1 MGELCRDSALTALDEETLWMMESHRRIRIVRCICPSRLTPYLROAKYLCQDDEEVVHSP 60  
Db 1 MGELCRDSALTALDEETLWMMESHRRIRIVRCICPSRLTPYLROAKYLCQDDEEVVHSP 60  
Qy 61 RLTSNAPRAGHLLDLKTRGNKGAIAFLSEIKFHNPDVYTLVTGLQPDVDFSNFSG--E 118  
Db 61 RLTSNAPRAGHLLDLKTRGNKGAIAFLSEIKFHNPDVYTLVTGLQPDVDFSNFSGME 120  
Qy 119 SSDF-DGLAGTSRNLRL 133  
Db 121 TSKLTCTCAGAGTSLQ 136

## RESULT 5

US-10-032-159a-18  
; Sequence 18, Application US/10032159A  
; Publication No. US20020164703A1  
; GENERAL INFORMATION:  
; APPLICANT: Pawlowski, Krzysztof  
; APPLICANT: Reed, John C.  
; TITLE OF INVENTION: CARD-DOMAIN CONTAINING POLYPEPTIDES,  
; FILE REFERENCE: P-LJ 5100  
; CURRENT APPLICATION NUMBER: US/10/032,159A  
; PRIOR FILING DATE: 2001-12-19  
; PRIOR APPLICATION NUMBER: US 60/257,457  
; PRIOR FILING DATE: 2000-12-21  
; NUMBER OF SEQ ID NOS: 37  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 18  
; LENGTH: 92  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-032-159a-18

Query Match 67.2%; Score 484; DB 14; Length 92;  
Best Local Similarity 100.0%; Pred. No. 2,2e-49;  
Matches 92; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 16 DEETLMMESHRRIRVRCICPSRLTPYLRQAKVLCOLDEEVLHSPRLTNSAMRAGHLL 75  
 DB 1 EETLMMESHRRIRVRCICPSRLTPYLRQAKVLCOLDEEVLHSPRLTNSAMRAGHLL 60  
 QY 76 DLKTRGNKGAIAFLBSLKFNPDVYTLVTGLQ 107  
 DB 61 LKTRGNKGAIAFLBSLKFNPDVYTLVTGLQ 92

RESULT 6  
 US-09-798-412-11  
 ; Sequence 11, Application US/09798412  
 ; Publication No. US20030109428A1

GENERAL INFORMATION:  
 APPLICANT: Berth, John  
 TITLE OF INVENTION: NOVEL MOLECULES OF THE CARD-RELATED  
 FILE REFERENCE: 07334-327001  
 CURRENT APPLICATION NUMBER: US/09/798,412  
 PRIOR FILING DATE: 2001-03-02  
 PRIOR APPLICATION NUMBER: US 09/728,260  
 PRIOR FILING DATE: 2000-12-01  
 PRIOR APPLICATION NUMBER: US 09/565,791  
 PRIOR FILING DATE: 2000-10-10  
 PRIOR APPLICATION NUMBER: US 09/513,904  
 PRIOR FILING DATE: 2000-02-25  
 PRIOR APPLICATION NUMBER: US 09/507,533  
 PRIOR FILING DATE: 2000-02-18  
 PRIOR APPLICATION NUMBER: US 60/168,780  
 PRIOR FILING DATE: 1999-12-03  
 NUMBER OF SEQ ID NOS: 19  
 SOFTWARE: FastSeq for Windows Version 4.0  
 SEQ ID NO 11  
 LENGTH: 1147  
 TYPE: PRT  
 ORGANISM: Homo sapiens  
 US-09-798-412-11

Query Match 37.6%; Score 270.5; DB 11; Length 1147;  
 Best Local Similarity 48.6%; Pred. No. 1.3e-22;  
 Matches 54; Conservative 23; Mismatches 33; Indels 1; Gaps 1;  
 QY 15 DEETLMMESHRRIRVRCICPSRLTPYLRQAKVLCOLDEEVLHSPRLTNSAMRAGHLL 74  
 DB 11 EEDALMENEVCNRMILSRYPINPAKLTPLYRQCKVIDEQDEDEVLANAPMLPSKINRAGRLL 70  
 QY 75 DLKTRGNKGAIAFLBSLKFNPDVYTLVTGLQPDVDFSNFSGSSDPDGL 125  
 DB 71 DLHTKGQGVVFLBSLEFYYPPELYKLTGKEPTRRPFSTIVVEEG-HEGL 120

RESULT 7  
 US-10-325-917-11  
 ; Sequence 11, Application US/10325917  
 ; Publication No. US20030113787A1  
 GENERAL INFORMATION:  
 APPLICANT: Berth, John  
 TITLE OF INVENTION: NOVEL MOLECULES OF THE CARD-RELATED  
 FILE REFERENCE: 07334-327001  
 CURRENT APPLICATION NUMBER: US/10/325,917  
 PRIOR FILING DATE: 2002-12-20  
 PRIOR APPLICATION NUMBER: US/09/798,412  
 PRIOR FILING DATE: 2001-03-02  
 PRIOR APPLICATION NUMBER: US 09/728,260  
 PRIOR FILING DATE: 2000-12-01  
 PRIOR APPLICATION NUMBER: US 09/565,791  
 PRIOR FILING DATE: 2000-10-10  
 PRIOR APPLICATION NUMBER: US 09/513,904  
 PRIOR FILING DATE: 2000-02-25  
 PRIOR APPLICATION NUMBER: US 09/507,533  
 PRIOR FILING DATE: 2000-02-18  
 PRIOR APPLICATION NUMBER: US 60/168,780

;; PRIOR FILING DATE: 1999-12-03  
 ;; NUMBER OF SEQ ID NOS: 19  
 ;; SOFTWARE: FastSeq for Windows Version 4.0  
 ;; SEQ ID NO 11  
 ;; LENGTH: 1147  
 ;; TYPE: PRT  
 ;; ORGANISM: Homo sapiens  
 US-10-325-917-11

Query Match 37.6%; Score 270.5; DB 15; Length 1147;  
 Best Local Similarity 48.6%; Pred. No. 1.3e-22;  
 Matches 54; Conservative 23; Mismatches 33; Indels 1; Gaps 1;

QY 15 DEETLMMESHRRIRVRCICPSRLTPYLRQAKVLCOLDEEVLHSPRLTNSAMRAGHLL 74  
 DB 11 EEDALMENEVCNRMILSRYPINPAKLTPLYRQCKVIDEQDEDEVLANAPMLPSKINRAGRLL 70  
 QY 75 DLKTRGNKGAIAFLBSLKFNPDVYTLVTGLQPDVDFSNFSGSSDPDGL 125  
 DB 71 DLHTKGQGVVFLBSLEFYYPPELYKLTGKEPTRRPFSTIVVEEG-HEGL 120

RESULT 8  
 US-10-032-159A-8  
 ; Sequence 8, Application US/10032159A  
 ; Publication No. US20020164703A1

GENERAL INFORMATION:  
 APPLICANT: Pawlowski, Krzysztof  
 APPLICANT: Reed, John C.  
 TITLE OF INVENTION: CARD-DOMAIN CONTAINING POLYPEPTIDES,  
 TITLE OF INVENTION: ENCODING NUCLEIC ACIDS, AND METHODS OF USE  
 FILE REFERENCE: P-1J 5100  
 CURRENT APPLICATION NUMBER: US/10/032,159A  
 PRIOR FILING DATE: 2001-12-19  
 PRIOR APPLICATION NUMBER: US 60/257,457  
 PRIOR FILING DATE: 2000-12-21  
 NUMBER OF SEQ ID NOS: 37  
 SOFTWARE: FastSeq for Windows Version 4.0  
 SEQ ID NO 8  
 LENGTH: 1247  
 TYPE: PRT  
 ORGANISM: Homo sapiens  
 US-10-032-159A-8

Query Match 37.6%; Score 270.5; DB 14; Length 1247;  
 Best Local Similarity 48.6%; Pred. No. 1.5e-22;  
 Matches 54; Conservative 23; Mismatches 33; Indels 1; Gaps 1;

QY 15 DEETLMMESHRRIRVRCICPSRLTPYLRQAKVLCOLDEEVLHSPRLTNSAMRAGHLL 74  
 DB 11 EEDALMENEVCNRMILSRYPINPAKLTPLYRQCKVIDEQDEDEVLANAPMLPSKINRAGRLL 70  
 QY 75 DLKTRGNKGAIAFLBSLKFNPDVYTLVTGLQPDVDFSNFSGSSDPDGL 125  
 DB 71 DLHTKGQGVVFLBSLEFYYPPELYKLTGKEPTRRPFSTIVVEEG-HEGL 120

RESULT 9  
 US-10-032-159A-10  
 ; Sequence 10, Application US/10032159A  
 ; Publication No. US20020164703A1  
 GENERAL INFORMATION:  
 APPLICANT: Pawlowski, Krzysztof  
 APPLICANT: Reed, John C.  
 APPLICANT: Godzik, Adam  
 TITLE OF INVENTION: CARD-DOMAIN CONTAINING POLYPEPTIDES,  
 TITLE OF INVENTION: ENCODING NUCLEIC ACIDS, AND METHODS OF USE  
 FILE REFERENCE: P-1J 5100  
 CURRENT APPLICATION NUMBER: US/10/032,159A  
 PRIOR FILING DATE: 2001-12-19  
 PRIOR APPLICATION NUMBER: US 60/257,457  
 PRIOR FILING DATE: 2000-12-21

NUMBER OF SEQ ID NOS: 37  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 10  
LENGTH: 92  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-10-032-159a-10

Query Match 35.7%; Score 257; DB 14; Length 92;  
Best Local Similarity 53.3%; Pred. No. 1.6e-22;  
Matches 48; Conservative 20; Mismatches 22; Indels 0; Gaps 0;

QY 16 DEETLWEMESHRRHIVRCICPSRLTPYLRQAKVLCQDDEEVLSHSPRLTNSAMRAGHLL 75  
DB 1 EDALWENECRHRHLSTRINPAKLTPLYRQCKVIDEDEDVLANAPLPSKINRAGRLLD 60  
QY 76 DLKTRGNKAIAFLSLKFNPDVYTLVTG 105  
DB 61 ILHTKGRGVVFFLESLEFFYPPELYKLVGTG 90

RESULT 10  
US-09-798-412-8  
Sequence 8, Application US/09798412  
Publication No. US20030109428A1  
GENERAL INFORMATION:  
APPLICANT: Berlin, John  
TITLE OF INVENTION: NOVEL MOLECULES OF THE CARD-RELATED  
FILE REFERENCE: 07334-327001  
CURRENT FILING DATE: US/09/798,412  
PRIOR FILING DATE: 2001-03-02  
PRIOR APPLICATION NUMBER: US 09/728,260  
PRIOR FILING DATE: 2000-12-01  
PRIOR APPLICATION NUMBER: US 09/685,791  
PRIOR FILING DATE: 2000-10-10  
PRIOR APPLICATION NUMBER: US 09/513,904  
PRIOR FILING DATE: 2000-02-25  
PRIOR APPLICATION NUMBER: US 09/507,533  
PRIOR FILING DATE: 2000-02-18  
PRIOR APPLICATION NUMBER: US 60/168,780  
PRIOR FILING DATE: 1999-12-03  
NUMBER OF SEQ ID NOS: 19  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 8  
LENGTH: 1032  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-09-798-412-8

Query Match 34.2%; Score 246; DB 11; Length 1032;  
Best Local Similarity 48.9%; Pred. No. 9.2e-20;  
Matches 46; Conservative 21; Mismatches 27; Indels 0; Gaps 0;

QY 15 DEETLWEMESHRRHIVRCICPSRLTPYLRQAKVLCQDDEEVLSHSPRLTNSAMRAGHLL 74  
DB 23 EDALWENECRHRHLSTRINPAKLTPLYRQCKVIDEDEDVLANAPLPSKINRAGRLLD 82  
QY 75 DLKTRGNKAIAFLSLKFNPDVYTLVTG 108  
DB 83 DLRCRGRGVVFFLESLEFFYPPEHFTLLTGOEP 116

RESULT 11  
US-10-325-917-8  
Sequence 8, Application US/10325917  
Publication No. US20030113787A1  
GENERAL INFORMATION:  
APPLICANT: Berlin, John  
TITLE OF INVENTION: NOVEL MOLECULES OF THE CARD-RELATED  
FILE REFERENCE: 07334-327001  
CURRENT APPLICATION NUMBER: US/10/325,917

CURRENT FILING DATE: 2002-12-20  
PRIOR APPLICATION NUMBER: US/09/798,412  
PRIOR FILING DATE: 2001-03-02  
PRIOR APPLICATION NUMBER: US 09/728,260  
PRIOR FILING DATE: 2000-12-01  
PRIOR APPLICATION NUMBER: US 09/685,791  
PRIOR FILING DATE: 2000-10-10  
PRIOR APPLICATION NUMBER: US 09/513,904  
PRIOR FILING DATE: 2000-02-25  
PRIOR APPLICATION NUMBER: US 09/507,533  
PRIOR FILING DATE: 2000-02-18  
PRIOR APPLICATION NUMBER: US 60/168,780  
NUMBER OF SEQ ID NOS: 19  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 8  
LENGTH: 1032  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-10-325-917-8

Query Match 34.2%; Score 246; DB 15; Length 1032;  
Best Local Similarity 48.9%; Pred. No. 9.2e-20;  
Matches 46; Conservative 21; Mismatches 27; Indels 0; Gaps 0;

QY 15 DEETLWEMESHRRHIVRCICPSRLTPYLRQAKVLCQDDEEVLSHSPRLTNSAMRAGHLL 74  
DB 23 EDALWENECRHRHLSTRINPAKLTPLYRQCKVIDEDEDVLANAPLPSKINRAGRLLD 82  
QY 75 DLKTRGNKAIAFLSLKFNPDVYTLVTG 108  
DB 83 DLRCRGRGVVFFLESLEFFYPPEHFTLLTGOEP 116

RESULT 12  
US-09-798-412-2  
Sequence 2, Application US/09798412  
Publication No. US20030109428A1  
GENERAL INFORMATION:  
APPLICANT: Berlin, John  
TITLE OF INVENTION: NOVEL MOLECULES OF THE CARD-RELATED  
FILE REFERENCE: 07334-327001  
CURRENT FILING DATE: US/09/798,412  
PRIOR FILING DATE: 2001-03-02  
PRIOR APPLICATION NUMBER: US 09/728,260  
PRIOR FILING DATE: 2000-12-01  
PRIOR APPLICATION NUMBER: US 09/685,791  
PRIOR FILING DATE: 2000-10-10  
PRIOR APPLICATION NUMBER: US 09/513,904  
PRIOR FILING DATE: 2000-02-25  
PRIOR APPLICATION NUMBER: US 09/507,533  
PRIOR FILING DATE: 2000-02-18  
PRIOR APPLICATION NUMBER: US 60/168,780  
NUMBER OF SEQ ID NOS: 19  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 2  
LENGTH: 536  
TYPE: PRT  
ORGANISM: Rattus norvegicus  
US-09-798-412-2

Query Match 31.0%; Score 223.5; DB 11; Length 536;  
Best Local Similarity 45.4%; Pred. No. 1.7e-17;  
Matches 49; Conservative 18; Mismatches 38; Indels 3; Gaps 1;

QY 15 DEETLWEMESHRRHIVRCICPSRLTPYLRQAKVLCQDDEEVLSHSPRLTNSAMRAGHLL 74  
DB 6 NDECSALESFRVKLISVIDPSRITPLYRQCKVLANPDDEEVLSPNLIKRVKGVLL 65  
QY 75 DLKTRGNKAIAFLSLKFNPDVYTLVTG 108  
DB 83 DLRCRGRGVVFFLESLEFFYPPEHFTLLTGOEP 116

Db 66 DILQRTGHKGVAFLFSLLELYPOLYRKVTGKEPARVFSMTIDAGS 113

```
RESULT 13
US-10-325-917-2
; Sequence 2, Application US/10325917
; Publication No. US20030113787A1
; GENERAL INFORMATION:
; APPLICANT: Bertin, John
; TITLE OF INVENTION: NOVEL MOLECULES OF THE CARD-RELATED
; FILE REFERENCE: 07334-327001
; CURRENT APPLICATION NUMBER: US/10/325,917
; PRIOR FILING DATE: 2002-12-20
; PRIOR APPLICATION NUMBER: US/09/798,412
; PRIOR FILING DATE: 2001-03-02
; PRIOR APPLICATION NUMBER: US 09/728,260
; PRIOR FILING DATE: 2000-12-01
; PRIOR APPLICATION NUMBER: US 09/685,791
; PRIOR FILING DATE: 2000-10-10
; PRIOR APPLICATION NUMBER: US 09/513,904
; PRIOR FILING DATE: 2000-02-25
; PRIOR APPLICATION NUMBER: US 09/507,533
; PRIOR FILING DATE: 2000-02-18
; PRIOR APPLICATION NUMBER: US 60/168,780
; PRIOR FILING DATE: 1999-12-03
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 536
; TYPE: PRT
; ORGANISM: Rattus norvegicus
US-10-325-917-2
```

```
Query Match 31.0%; Score 223.5; DB 15; Length 536;
Best Local Similarity 45.4%; Pred. No. 1.7e-17;
Matches 49; Conservative 18; Mismatches 38; Indels 3; Gaps 1;

QY 15 DEETLMMESHRRHRIYRCICPSRLTPYLRQAKVLCQDDEEVLHSPRLTNSAMRAGHLL 74
:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 6 NDECMWVLEGFRTVLTLSVIDPSRITPYLRQCKVLPDDEQVLSDPNLTIRKRVGVLL 65

QY 75 DLKTRGNKGAIAFLFSLKFNPDVYTLVTGLQPDVDS--NFSGES 119
:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 66 DILQRTGHKGVAFLFSLLELYPOLYRKVTGKEPARVFSMTIDAGS 113
```

```
RESULT 14
US-10-032-159A-2
; Sequence 2, Application US/10032159A
; Publication No. US20020164703A1
; GENERAL INFORMATION:
; APPLICANT: Pawlowski, Krzysztof
; APPLICANT: Reed, John C.
; APPLICANT: Godzik, Adam
; TITLE OF INVENTION: CARD-DOMAIN CONTAINING POLYPEPTIDES,
; FILE REFERENCE: P-LJ 5100
; CURRENT APPLICATION NUMBER: US/10/032,159A
; PRIOR FILING DATE: 2001-12-19
; PRIOR APPLICATION NUMBER: US 60/257,457
; PRIOR FILING DATE: 2000-12-21
; NUMBER OF SEQ ID NOS: 37
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 366
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-032-159A-2
```

```
Query Match 29.9%; Score 215.5; DB 14; Length 366;
Best Local Similarity 44.4%; Pred. No. 8.8e-17;
Matches 48; Conservative 17; Mismatches 40; Indels 3; Gaps 1;
```

```
QY 15 DEETLMMESHRRHRIYRCICPSRLTPYLRQAKVLCQDDEEVLHSPRLTNSAMRAGHLL 74
:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 6 NDECMWVLEGFRTVLTLSVIDPSRITPYLRQCKVLPDDEQVLSDPNLTIRKRVGVLL 65

QY 75 DLKTRGNKGAIAFLFSLKFNPDVYTLVTGLQPDVDS--NFSGES 119
:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 66 DILQRTGHKGVAFLFSLLELYPOLYRKVTGKEPARVFSMTIDAGS 113
```

```
RESULT 15
US-10-032-159A-20
; Sequence 20, Application US/10032159A
; Publication No. US20020164703A1
; GENERAL INFORMATION:
; APPLICANT: Pawlowski, Krzysztof
; APPLICANT: Reed, John C.
; APPLICANT: Godzik, Adam
; TITLE OF INVENTION: CARD-DOMAIN CONTAINING POLYPEPTIDES,
; FILE REFERENCE: P-LJ 5100
; CURRENT APPLICATION NUMBER: US/10/032,159A
; PRIOR FILING DATE: 2001-12-19
; PRIOR APPLICATION NUMBER: US 60/257,457
; PRIOR FILING DATE: 2000-12-21
; NUMBER OF SEQ ID NOS: 37
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 20
; LENGTH: 366
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-032-159A-20
```

```
Query Match 29.9%; Score 215.5; DB 14; Length 366;
Best Local Similarity 44.4%; Pred. No. 8.8e-17;
Matches 48; Conservative 17; Mismatches 40; Indels 3; Gaps 1;

QY 15 DEETLMMESHRRHRIYRCICPSRLTPYLRQAKVLCQDDEEVLHSPRLTNSAMRAGHLL 74
:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 6 NDECMWVLEGFRTVLTLSVIDPSRITPYLRQCKVLPDDEQVLSDPNLTIRKRVGVLL 65

QY 75 DLKTRGNKGAIAFLFSLKFNPDVYTLVTGLQPDVDS--NFSGES 119
:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 66 DILQRTGHKGVAFLFSLLELYPOLYRKVTGKEPARVFSMTIDAGS 113
```

Search completed: February 18, 2004, 04:56:50  
Job time : 78 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

## OM nucleic - nucleic search, using sw model

Run on: February 17, 2004, 21:36:16 ; Search time 1366.86 Seconds  
(without alignments)  
8260.598 Million cell updates/sec

Title: US-10-032-159A-17

Perfect score: 276  
Sequence: 1 gagagagacactgtgagagat.....ccctgtcacccggcgtagcag 276

Scoring table:  
IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 2888711 seqs, 2045813386 residues

Total number of hits satisfying chosen parameters: 5777422

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

## Database :

GenEmbl:\*  
1: gb\_ba:\*  
2: gb\_hng:\*  
3: gb\_in:\*  
4: gb\_om:\*  
5: gb\_ov:\*  
6: gb\_pat:\*  
7: gb\_ph:\*  
8: gb\_pl:\*  
9: gb\_pr:\*  
10: gb\_ro:\*  
11: gb\_sts:\*  
12: gb\_sy:\*  
13: gb\_un:\*  
14: gb\_vt:\*  
15: em\_ba:\*  
16: em\_fun:\*  
17: em\_hum:\*  
18: em\_in:\*  
19: em\_mu:\*  
20: em\_om:\*  
21: em\_or:\*  
22: em\_ov:\*  
23: em\_pat:\*  
24: em\_ph:\*  
25: em\_pl:\*  
26: em\_ro:\*  
27: em\_sts:\*  
28: em\_un:\*  
29: em\_vt:\*  
30: em\_hng\_hum:\*  
31: em\_hng\_inv:\*  
32: em\_hng\_other:\*  
33: em\_hng\_mus:\*  
34: em\_hng\_pln:\*  
35: em\_hng\_rnd:\*  
36: em\_hng\_mam:\*  
37: em\_hng\_vrt:\*  
38: em\_sy:\*  
39: em\_hngo\_hum:\*  
40: em\_hngo\_mus:\*  
41: em\_hngo\_other:\*

score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	276	100.0	2621	9 BC018142	BC018142 Homo sapi
2	276	100.0	3682	9 AY032927	AY032927 Homo sapi
3	276	100.0	3766	9 AK091123	AK091123 Homo sapi
4	276	100.0	3951	9 AF322642	AF322642 Homo sapi
5	213.6	77.4	3772	10 AF363457	AF363457 Mus muscu
6	213.6	77.4	3995	10 BC029102	BC029102 Mus muscu
7	167	60.5	55173	2 AC132816	AC132816 Homo sapi
8	167	60.5	144000	9 AC123764	AC123764 Homo sapi
9	167	60.5	182016	9 AC087741	AC087741 Homo sapi
10	167	60.5	187865	2 AC015559	AC015559 Homo sapi
11	167	60.5	207822	2 AC109321	AC109321 Homo sapi
12	143.6	52.0	74301	2 AC137736	AC137736 Homo sapi
13	131.6	47.7	65608	2 AC132197	AC132197 Homo sapi
14	122.4	44.3	24653	10 AL645811	AL645811 Mouse DNA
15	113.4	41.1	3096	6 AX154570	AX154570 Sequence
16	113.4	41.1	3908	9 AY032928	AY032928 Homo sapi
17	113.4	41.1	3948	6 AX300499	AX300499 Sequence
18	113.4	41.1	3949	6 AX154568	AX154568 Sequence
19	113.4	41.1	3949	9 AY028896	AY028896 Homo sapi
20	111	40.2	55173	2 AC132816	AC132816 Homo sapi
21	110.6	40.1	4085	10 AY135367	AY135367 Mus muscu
22	105.4	38.2	4438	10 AF363456	AF363456 Mus muscu
23	100.4	36.4	3441	6 AX154573	AX154573 Sequence
24	100.4	36.4	3955	9 AF352576	AF352576 Homo sapi
25	100.4	36.4	4276	6 AX154571	AX154571 Sequence
26	100.4	36.4	4276	6 AF322641	AF322641 Homo sapi
27	100.4	36.4	4293	9 AK074049	AK074049 Homo sapi
28	88	31.9	765	6 BD150612	BD150612 Primer fo
29	88	31.9	1608	6 AX154567	AX154567 Sequence
30	88	31.9	2098	6 AX154565	AX154565 Sequence
31	88	31.9	2132	9 AF311287	AF311287 Homo sapi
32	88	31.9	2176	6 BD160313	BD160313 Primer fo
33	88	31.9	2176	6 AK024001	AK024001 Homo sapi
34	86.4	31.3	1836	9 BC008877	BC008877 Homo sapi
35	79.6	28.8	1608	6 AX154564	AX154564 Sequence
36	79.6	28.8	1879	6 AX154562	AX154562 Sequence
37	79.6	28.8	1879	10 AF311288	AF311288 Rattus no
38	72.6	26.3	239392	2 AC115417	AC115417 Rattus no
39	72.6	26.3	245032	2 AC109749	AC109749 Rattus no
40	72.4	26.2	96256	9 HS117715	HS117715 Human DNA
41	69.4	25.1	187446	2 AC124374	AC124374 Mus muscu
42	69.4	25.1	225027	2 AC131725	AC131725 Mus muscu
43	68.8	24.9	3164	10 BC004692	BC004692 Mus muscu
44	67.8	24.6	161985	10 AL592169	AL592169 Mouse DNA
45	67.8	24.6	217819	2 AC026386	AC026386 Mus muscu

## ALIGNMENTS

RESULT 1  
BC018142  
LOCUS 2621 bp mRNA 1linear PRI 06-DEC-2001  
DEFINITION Homo sapiens, Similar to caspase recruitment domain protein 14,  
clone MGC:9539 IMAGE:3847282, mRNA, complete cds.  
ACCESSION BC018142  
VERSION BC018142.1 GI:17390314  
KEYWORDS MGC.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1 (bases 1 to 2621)  
AUTHORS Strausberg, R.  
TITLE Direct Submission

## JOURNAL

Submitted (03-DEC-2001) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA

REMARK  
COMMENT

NIH-MGC Project URL: <http://mgc.nci.nih.gov>  
Contact: MGC help desk  
Email: [cgaps-remail.nih.gov](mailto:cgaps-remail.nih.gov)  
Tissue Procurement: ATCC  
DNA Library Preparation: Life Technologies, Inc.  
DNA Sequencing by: The I.M.A.G.E. Consortium (LNL)  
Center, Stanford University School of Medicine, Stanford, CA 94305  
Web site: <http://www-shgc.stanford.edu>  
Contact: (Dickson, Mark) [mcdpaxil.stanford.edu](mailto:mcdpaxil.stanford.edu)  
Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers, R. M.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at: <http://image.lnl.gov>  
Series: IRAC plates: 20 Row: 1 Column: 3  
This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 16507953.

FEATURES  
source

1. 2621  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="MGC:9539 IMAGE:3847282"  
/tissue\_type="Colon, adenocarcinoma"  
/clone\_id="NIH MGC\_65"  
/lab\_host="DH10B"  
/note="Vector: pCMV-SPORT6"  
178..2400  
/codon\_start=1  
/product="Similar to caspase recruitment domain protein 14"  
/protein\_id="AAH18142.1"  
/db\_xref="GI:17380315"  
/translation="MSELCRDSALVATLDEETLMEWESHRRIVRCICPSRLTPYLQAVYLQDLDEEVSLSPRLTNSAMRAGHLIDLKTRGKGAIAFLSLFHPNDVTLVTGLQPDVDFNSFGLMETSKLTCGLAALISLOELNOKGQKVELLRQCOJELHGLAETRAEGHOLEADHSRMKREVAHFHEVLRKDEMLSLSHYSNALOEKLAARCRSLQELYLILKQELQKRNAYVSSCELEHQOSLRITASDQSGDEBELNKEBEXKRS LTFSLAEKDILOSLEDEARSRQELVERHSLERAVAEQREQYWEKEQTLQFO KSKRAQOLYREKYNALQAVCELOKXERDQYASRDSQRELSQILVEKQSLRQVEL TDVCELTQRLQLOAPPGVILKQEARTRPCPEKQRLVYMAHICPRDSDCLSVS TESQSLDLSATSRRELVDSPRSSPAPSOQSLYKRAVEDPGEPPMFSQCEIPEG DPGALPKAKGDDPHLDYELDPLADLPOLSSLOPVSPGRIDVSEGVLMRRPARIL SQVTMLAFQGDALBEQISVTGMDTGLFIHRTVPGSLADQMLRPGQIVWDYVASE PLFRAVLDDTLLEAVGLLRVVDGFCCLSVKVTIDGKRLQDLKAVATSGDSFYIR VNLAHEGRAGELOVHCNEVILVHTDTFCGCGCMHARVNSYTMKDTLAHGTTPNYSR

## CDS

BASE COUNT 563 a 782 c 857 g 419 t  
ORIGIN

Query Match 100.0%; Score 276; DB 9; Length 2621;  
Best Local Similarity 100.0%; Pred. No. 2.9e-44;  
Matches 276; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GAGAGACACTGTGGAGATGATGAGAGACCGCCACAGATCGTACCTGATTCG 60  
DB 223 GAGAGACACTGTGGAGATGATGAGAGACCGCCACAGATCGTACCTGATTCG 282  
QY 61 CCCAGCGGCTCACCCCTACCTGCGCAGGCGCAAGTGTGCGCAGCTGACGAGAG 120  
DB 283 CCCAGCGGCTCACCCCTACCTGCGCAGGCGCAAGTGTGCGCAGCTGACGAGAG 342  
QY 121 GAGGTGCTGACAGCCCGGCTCACCAACAGCCATGCGGCGCGGCACTTGCTGAT 180  
DB 343 GAGGTGCTGACAGCCCGGCTCACCAACAGCCATGCGGCGCGGCACTTGCTGAT 402  
QY 181 TTGCTGAAGACTCGAGGAGAGACGGGCGCATCGCTTCTGAGAGACCTGAAGTTCCAC 240

## DB

403 TTGCTGAAGACTCGAGGAGAGACGGGCGCATCGCTTCTGAGAGACCTGAAGTTCCAC 462  
QY 241 AACCTGACGTCTACACCCCTGATCACCGGGCTGACG 276  
DB 463 AACCTGACGTCTACACCCCTGATCACCGGGCTGACG 498

RESULT 2  
AY032927  
LOCUS  
DEFINITION  
Homo sapiens CARD-containing MAGUK 2 protein (CARMA2) mRNA,  
complete cds.  
ACCESSION  
AY032927  
VERSION  
AY032927.1 GI:14192722  
KEYWORDS  
SOURCE  
ORGANISM  
Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
1 (bases 1 to 3682)  
Gaido, O., Martignon, F., Mischeu, O., Bonnet, D., Thome, M. and  
Tschopp, J.  
Carnal, a CARD-containing binding partner of Bcl10, induces Bcl10  
phosphorylation and NF-kappaB activation  
FEBS Lett. 496 (2-3), 121-127 (2001)  
11356195  
2 (bases 1 to 3682)  
Martignon, F., Thome, M. and Tschopp, J.  
Direct Submission  
Submitted (23-Apr-2001) Institute of Biochemistry, University of  
Lausanne, Ch. des Boveresses 155, Epalinges, VD 1066, Switzerland  
Location/Qualifiers  
1..3682  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
1..3682  
/gene="CARMA2"  
4..3018  
/gene="CARMA2"  
/note="contains caspase recruitment domain CARD, coiled  
coil, PDZ, SH3, guanylate kinase domain GUK, interacts  
with the CARD-containing molecule Bcl10 (CARMEN)"  
/codon\_start=1  
/product="CARD-containing MAGUK 2 protein"  
/protein\_id="AAK54453.1"  
/db\_xref="GI:14192723"  
/translation="MSELCRDSALVATLDEETLMEWESHRRIVRCICPSRLTPYLQAVYLQDLDEEVSLSPRLTNSAMRAGHLIDLKTRGKGAIAFLSLFHPNDVTLVTGLQPDVDFNSFGLMETSKLTCGLAALISLOELNOKGQKVELLRQCOJELHGLAETRAEGHOLEADHSRMKREVAHFHEVLRKDEMLSLSHYSNALOEKLAARCRSLQELYLILKQELQKRNAYVSSCELEHQOSLRITASDQSGDEBELNKEBEXKRS LTFSLAEKDILOSLEDEARSRQELVERHSLERAVAEQREQYWEKEQTLQFO KSKRAQOLYREKYNALQAVCELOKXERDQYASRDSQRELSQILVEKQSLRQVEL TDVCELTQRLQLOAPPGVILKQEARTRPCPEKQRLVYMAHICPRDSDCLSVS TESQSLDLSATSRRELVDSPRSSPAPSOQSLYKRAVEDPGEPPMFSQCEIPEG DPGALPKAKGDDPHLDYELDPLADLPOLSSLOPVSPGRIDVSEGVLMRRPARIL SQVTMLAFQGDALBEQISVTGMDTGLFIHRTVPGSLADQMLRPGQIVWDYVASE PLFRAVLDDTLLEAVGLLRVVDGFCCLSVKVTIDGKRLQDLKAVATSGDSFYIR VNLAHEGRAGELOVHCNEVILVHTDTFCGCGCMHARVNSYTMKDTLAHGTTPNYSR

FEATURES  
source

gene  
CDS  
BASE COUNT 751 a 1100 c 1194 g 637 t  
ORIGIN  
Query Match 100.0%; Score 276; DB 9; Length 3682;  
Best Local Similarity 100.0%; Pred. No. 2.8e-44;  
Matches 276; Conservative 0; Mismatches 0; Indels 0; Gaps 0;